DECLARATION

- I, Akiko KOSEMURA, of HIRAKI & ASSOCIATES, do solemnly and sincerely declare as follows:
- That I am well acquainted with the English and Japanese languages and am competent to translate from Japanese into English.
- That I have executed, with the best of my ability, a true and correct translation into English of Japanese Patent Application No. 329115/2003 filed on September 19, 2003, a copy of which I attach herewith.

This 16th day of July, 2010

VICITY A COMPANIED V

[Title of Document] CLAIMS

[Claim 1]

A replicon RNA, comprising a nucleotide sequence containing at least the 5' untranslated region, the nucleotide sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein and the 3' untranslated region on the genomic RNA of hepatitis C virus of genotype 2a.

[Claim 2]

The replicon RNA of claim 1, containing at least one selection marker gene or a reporter gene, and at least one IRES sequence.

[Claim 3]

A replicon RNA, comprising a nucleotide sequence containing the 5' untranslated region comprising the nucleotide sequence represented by SEQ ID NO: 9 or 10; at least one selection marker gene or a reporter gene; an IRES sequence; the nucleotide sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein on the genomic RNA of hepatitis C virus of genotype 2a; and the 3' untranslated region comprising the nucleotide sequence represented by SEO ID NO: 11 or 12.

[Claim 4]

The replicon RNA of any one of claims 1 to 3, wherein the genomic RNA of hepatitis C virus of genotype 2a is an RNA comprising the nucleotide sequence represented by SEQ ID NO: 3 or 5.

[Claim 5]

A replicon RNA, comprising the following RNA (a) or (b):

- (a) an RNA comprising the nucleotide sequence represented by SEQ ID NO: 1 or 2; and
- (b) an RNA comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 or 2 by deletion, substitution or addition of 1 to 10 nucleotides, and being capable of autonomous replication.

[Claim 6]

A replicon-replicating cell, which is prepared by introducing the replicon RNA of any one of claims 1 to 5 into a cell.

[Claim 7]

The replican-replicating cell of claim 6, wherein the cell is a eukaryotic cell.

[Claim 8]

The replicon-replicating cell of claim 7, wherein the eukaryotic cell is a human liver-derived cell, a human uterine cervix-derived cell or a human fetal kidney-derived cell.

[Claim 9]

The replicon-replicating cell of claim 7, wherein the eukaryotic cell is any one cell selected from the group consisting of an Huh7 cell, an HepG2 cell, an IMY-N9 cell, an HeLa cell and a 293 cell.

[Claim 10]

The replicon RNA of any one of claims I to 5, which is for producing or evaluating a therapeutic agent or a diagnostic agent against hepatitis C virus infection.

[Claim 11]

The replican-replicating cell of any one of claims 6 to 9, which is for producing or evaluating a therapeutic agent or a diagnostic agent against hepatitis C virus infection.

[Claim 12]

The replicon RNA of any one of claims 1 to 5, which is for producing a vaccine against hepatitis C virus infection.

[Claim 13]

The replicon-replicating cell of any one of claims 6 to 9, which is for producing a vaccine against hepatitis C virus infection.

[Claim 14]

A method of producing a replicon RNA of hepatitis C virus of genotype 2a, comprising extracting the replicon RNA from the replicon-replicating cell of any

one of claims 6 to 9.

[Claim 15]

A method of producing a viral protein of hepatitis C virus of genotype 2a, comprising culturing the replicon-replicating cell of any one of claims 6 to 9, and obtaining the viral protein from the resulting culture product.

[Claim 16]

A method of screening for a substance promoting or suppressing the replication of hepatitis C virus, comprising culturing the replicon-replicating cell of any one of claims 6 to 9 in the presence of a test substance, and detecting the replication of a replicon RNA in the resulting culture product.

[Claim 17]

A method of increasing the replication efficiency of the replicon RNA of hepatitis C virus of genotype 2a, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of any one of claims 6 to 9, and introducing the thus obtained replicated replicon RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell.

[Claim 18]

The method of claim 17, wherein the replication efficiency increases to become at least two times greater than that of the replicon RNA that is introduced at the beginning into the replicon-replicating cell.

(Claim 191

A method of producing a replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of any one of claims 6 to 9, and introducing the thus obtained replicated replicon RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell; and obtaining a replicated replicon RNA from the finally obtained replicon-replicating cell.

[Claim 20]

A method of producing a replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency, comprising detecting a nucleotide mutation or an amino acid mutation between the replicon RNA that is produced so as to have an increased replication efficiency by the method of claim 19 and the replicon RNA that is introduced at the beginning into the replicon-replicating cell; and introducing the thus detected nucleotide mutation or amino acid mutation into a replicon RNA whose replication efficiency is to be increased.

(Claim 211

A replicon RNA, comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 by at least one mutation selected from the group consisting of the following (a) to (u):

- (a) a mutation from A to G at nucleotide site 7157;
- (b) a mutation from C to U at nucleotide site 4955;
- (c) a mutation from A to G at nucleotide site 4936;
- (d) a mutation from A to G at nucleotide site 5000;
- (e) a mutation from A to G at nucleotide site 7288;
- (f) a mutation from G to U at nucleotide site 5901;
- (g) a mutation from A to U at nucleotide site 6113;
- (h) a mutation from A to G at nucleotide site 2890;
- (i) a mutation from C to A at nucleotide site 6826;
- (i) a mutation from C to A at nucleotide site 6887;
- ---
- (k) a mutation from U to A at nucleotide site 6580;
- (1) a mutation from U to C at nucleotide site 7159;
- (m) a mutation from U to A at nucleotide site 7230;
- (n) a mutation from C to A at nucleotide site 6943;
- (o) a mutation from G to A at nucleotide site 5687;
- (p) a mutation from A to G at nucleotide site 6110;
- (g) a mutation from U to C at nucleotide site 5550;

- (r) a mutation from A to G at nucleotide site 7217;
- (s) a mutation from A to G at nucleotide site 3643;
- (t) a mutation from G to A at nucleotide site 5851; and
- (u) a mutation from G to A at nucleotide site 5914.

[Title of Document] DESCRIPTION

[Title Of Invention] A NUCLEIC ACID CONSTRUCT CONTAINING A NUCLEIC ACID DERIVED FROM THE GENOME OF HEPATITIS C VIRUS (HCV) OF GENOTYPE 2a, AND A CELL HAVING SUCH NUCLEIC ACID CONSTRUCT INTRODUCED THEREIN

[Technical Field]

[0001]

The present invention relates to a replicon RNA of the hepatitis C virus of genotype 2a, a replicon-replicating cell wherein the replicon RNA is introduced, and a method of increasing the replication efficiency of the replicon RNA.

[Background Art]

[0002]

The hepatitis C virus (HCV) is a virus belonging to the family Flaviviridae. It has a single-stranded (+) strand sense RNA as its genome and is known to cause hepatitis C. Recent studies have revealed that Hepatitis C virus is classified into a number of types based on genotypes or serotypes. According to the phylogenetic analysis of Simmonds et al., using the nucleotide sequences of the HCV strains, which is currently a mainstream method of classifying HCV genotypes, HCV is classified into 6 genotypes: genotype 1a, genotype 1b, genotype 2a, genotype 2b, genotype 3a and genotype 3b (see Non Patent Literature 1). Each of these types is further classified into several subtypes. The nucleotide sequences of the full-length genomes of a several number of genotypes of HCV have been determined to date (see Patent Literature 1 and Non Patent Literatures 2-5).

[0003]

HCV causes chronic hepatitis by persistent infection. Currently, the main cause of chronic hepatitis observed worldwide is persistent HCV infection.

Actually, around 50% of individuals with persistent infection develop chronic hepatitis. Chronic hepatitis in approximately 20% of these patients shifts to liver

cirrhosis over the course of 10 to 20 years, and some of these patients further go on to advanced lethal pathological conditions such as hepatic cancer.

Hepatitis C is currently treated mainly by a therapy using interferon- α or interferon- β , or a therapy using in combination interferon- α and ribavirin, the purine-nucleoside derivative. However, even when these therapies are performed, the therapeutic effects are observed in only approximately 60% of all the treated patients. When the therapies are ceased after the exertion of the effects, the disease recrudesces in more than half of the patients. The therapeutic effect of interferones is known to relate to HCV genotypes, and is said to be lower against genotype 1b and higher against genotype 2a (see Non Patent Literature 6).

It is an important goal to develop therapeutic agents or prophylactic agents effective against hepatitis C, the incidence rate of which is high in industrial countries, for which currently no causal treatment are present, and which finally bring about serious results. Hence, the development of HCV-specific chemotherapies and vaccine therapies are earnestly desired. A target for the development of an anti-HCV agent may be the suppression of HCV replication or the suppression of infection of cells with HCV.

Until recently, propagation of HCV in a cell culture system and infecting cultured cells with HCV have been difficult. Moreover, a chimpanzee has been the only animal that can be infected with HCV and can be used in experiments, so that it has been difficult to carry out studies on the replication mechanism of HCV and the infection mechanism of HCV. However, recently, HCV subgenomic RNA replicons have been prepared as HCV-derived autonomously replicable RNA (see Patent Literature 2 and Non Patent Literatures 7-10), which enables the analysis of the replication mechanism of HCV using cultured cells. These HCV subgenomic RNA replicons are each prepared by substituting structural proteins existing

downstream of HCV IRES in the 5' untranslated region of the HCV genomic RNA of genotype 1b with a neomycin resistance gene and EMCV IRES that has been ligated downstream of the resistance gene. It has been demonstrated that this RNA replicon is autonomously replicated in human hepatic cancer cells, Huh7 cells, when introduced into the Huh7 cells followed by culture in the presence of neomycin.

[0007]

However, regarding such intracellular RNA replication systems for HCV, only those using HCV genomic RNA of genotype Ib have been prepared so far. Since there has been a report that different genotypes of HCV differ also in viral proteins encoded, it may be difficult to sufficiently elucidate the replication mechanism of HCV only by analyzing the subgenomic RNA replicons derived from HCV of genotype 1b. Furthermore, based on the fact that the therapeutic effects of interferons differ depending on the HCV genotypes, it may be particularly difficult to develop an anti-HCV agent having an effect on various types of HCV by the use of only an HCV replication system containing the subgenomic RNA replicon of HCV of genotype 1b.

[8000]

(1992) 183, pp. 334-342

[Patent Literature 1] JP Patent Publication (Kokai) No. 2002-171978 A

[Patent Literature 2] JP Patent Publication (Kokai) No. 2001-17187 A

[Non Patent Literature 1] Simmonds, P. et al, Hepatology, (1994) 10, pp. 1321-

[Non Patent Literature 2] Choo et al., Science, (1989) 244, pp. 359-362 [Non Patent Literature 3] Kato et al., J. Med. Virol., (2001) 64(3) pp. 334-339 [Non Patent Literature 4] Okamoto, H et al, J. Gen. Virol., (1992) 73 pp. 673-679 [Non Patent Literature 5] Mori, S. et al, Biochem. Biophis. Res. Commun.,

[Non Patent Literature 6] Yoshioka et al., Hepatology, (1992) 16(2): pp. 293-299 [Non Patent Literature 7] Lohmann et al., Science, (1999) 285, pp. 110-113 [Non Patent Literature 8] Blight et al., Science, (2000) 290, pp. 1972-1974
[Non Patent Literature 9] Friebe et al., J. Virol., (2001) 75(24): pp. 12047-12057
[Non Patent Literature 10] Ikeda et al., J. Virol., (2002) 76(6): pp. 2997-3006
[Disclosure of Invention]
[Problem to be Solved by Invention]

An object of the present invention is to provide an HCV-derived replicon RNA of a HCV genotype for which replicon RNA has not yet been prepared.

[Means for Solving the Problem]

[0100]

[0009]

As a result of intensive studies to achieve the above object, we have succeeded in preparing the replicon RNA of HCV genotype 2a. [0011]

That is, the present invention is as follows.

- [1] A replicon RNA, comprising a nucleotide sequence containing at least the 5' untranslated region, the nucleotide sequence encoding NS3 protein, NS4A protein, NS5B protein and NS5B protein and the 3' untranslated region on the genomic RNA of hepatitis C virus of genotype 2a. Preferably, this replicon RNA further contains at least one selection marker gene or a reporter gene, and at least one IRES sequence.
- [2] A replicon RNA, comprising a nucleotide sequence containing the 5' untranslated region comprising the nucleotide sequence represented by either SEQ ID NO: 9 or 10; at least one selection marker gene or a reporter gene; an IRES sequence; the nucleotide sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein on the genomic RNA of hepatitis C virus of genotype 2a; and the 3' untranslated region comprising the nucleotide sequence represented by either SEQ ID NO: 11 or 12.
- [3] The replicon RNA of [1] or [2] above, wherein the genomic RNA of hepatitis C virus of genotype 2a is an RNA comprising the nucleotide sequence represented by

- SEO ID NO: 3 or 5.
- [4] A replicon RNA, comprising the following RNA (a) or (b):
- (a) an RNA comprising the nucleotide sequence represented by SEQ ID NO: 1 or 2: and
- (b) an RNA comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 or 2 by deletion, substitution or addition of 1 to 10 nucleotides, and being capable of autonomous replication.
- [5] A replicon-replicating cell, which is prepared by introducing the replicon RNA of any one of [1] to [4] above into a cell. For this replicon-replicating cell, a cell into which the replicon RNA is introduced is preferably a cukaryotic cell, more preferably a human liver-derived cell, a human uterine cervix-derived cell or a human fetal kidney-derived cell, and further more preferably any one cell selected from the group consisting of an Huh7 cell, an HepG2 cell, an IMY-N9 cell, an HeLa cell and a 293 cell.
- [6] The replicon RNA of [1] to [4] above, which is for producing or evaluating a therapeutic agent or a diagnostic agent against hepatitis C virus infection.
- [7] The replicon-replicating cell of [5] above, which is for producing or evaluating a therapeutic agent or a diagnostic agent against hepatitis C virus infection.
- [8] The replicon RNA of [1] to [4] above, which is for producing a vaccine against hepatitis C virus infection.
- [9] The replicon-replicating cell of [5] above, which is for producing a vaccine against hepatitis C virus infection.
- [10] A method of producing a replicon RNA of hepatitis C virus of genotype 2a, comprising extracting the replicon RNA from the replicon-replicating cell of [5] above.
- [11] A method of producing a viral protein of hepatitis C virus of genotype 2a, comprising culturing the replicon-replicating cell of [5] above, and obtaining the viral protein from the resulting culture product.
- [12] A method of screening for a substance promoting or suppressing the

replication of hepatitis C virus, comprising culturing the replicon-replicating cell of [5] above in the presence of a test substance, and detecting the replication of a replicon RNA in the resulting culture product.

- [13] A method of increasing the replication efficiency of the replicon RNA of hepatitis C virus of genotype 2a, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of [5] above, and introducing the thus obtained replicated replicon RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell. In this method, it is more preferred that the replication efficiency increases to become preferably at least two times greater than that of the replicon RNA that is introduced at the beginning into the replicon-replicating cell.
- [14] A method of producing a replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency, comprising performing once or more the following: obtaining a replicated replicon RNA from the replicon-replicating cell of [5] above, and introducing the thus obtained replicated replicon RNA into a cell that is different from the replicon-replicating cell so as to prepare a new replicon-replicating cell; and obtaining a replicated replicon RNA from the finally obtained replicon-replicating cell.
- [15] A method of producing a replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency, comprising detecting a nucleotide mutation or an amino acid mutation between the replicon RNA that is produced so as to have an increased replication efficiency by the method of [14] above and the replicon RNA that is introduced at the beginning into the replicon-replicating cell; and introducing the thus detected nucleotide mutation or amino acid mutation into a replicon RNA whose replication efficiency is to be increased.
- [16] A replicon RNA, comprising a nucleotide sequence derived from the nucleotide sequence represented by SEQ ID NO: 1 by at least one mutation selected from the group consisting of the following (a) to (u):

- (a) a mutation from A to G at nucleotide site 7157;
- (b) a mutation from C to U at nucleotide site 4955;
- (c) a mutation from A to G at nucleotide site 4936;
- (d) a mutation from A to G at nucleotide site 5000;
- (e) a mutation from A to G at nucleotide site 7288;
- (f) a mutation from G to U at nucleotide site 5901;
- (g) a mutation from A to U at nucleotide site 6113;
- (h) a mutation from A to G at nucleotide site 2890;
- (i) a mutation from C to A at nucleotide site 6826;
- (i) a mutation from C to A at nucleotide site 6887;
- () a matation from a to 71 at matrices and a 111
- (k) a mutation from U to A at nucleotide site 6580;
- (I) a mutation from U to C at nucleotide site 7159;
- (m) a mutation from U to A at nucleotide site 7230;
- (n) a mutation from C to A at nucleotide site 6943;
- (o) a mutation from G to A at nucleotide site 5687;
- (p) a mutation from A to G at nucleotide site 6110;
- (a) a mutation from U to C at nucleotide site 5550;
- (r) a mutation from A to G at nucleotide site 7217;
- (s) a mutation from A to G at nucleotide site 3643;
- (t) a mutation from G to A at nucleotide site 5851; and
- (u) a mutation from G to A at nucleotide site 5914.

[Effects of Invention]

[0012]

According to the present invention, an HCV-RNA replicon derived from the genotype 2a strain of HCV has been provided for the first time. The repliconreplicating cell according to the present invention can be used as a culture system for the continuous production of RNA and HCV proteins derived from HCV of genotype 2a. Furthermore, the replicon-replicating cell according to the present invention is useful as a test system for screening for various substances that affect

HCV replication and/or the translation of HCV proteins.

[Best Mode for Carrying out Invention]

[0013]

The present invention is explained in detail as follows.

[0014]

1, HCV-derived replicon RNA according to the present invention

The genome of hepatitis C virus (HCV) is a single-stranded (+) strand RNA comprising approximately 9600 nucleotides. This genomic RNA comprises the 5' untranslated region (also denoted as 5' NTR or 5' UTR), a translated region composed of a structural region and a non-structural region and the 3' untranslated region (also denoted as 3' NTR or 3' UTR). HCV structural proteins are encoded in the structural region, and a plurality of non-structural proteins are encoded in the non-structural region.

[0015]

Such HCV structural proteins and non-structural proteins are generated through the translation into a continuous form thereof, a polyprotein, from the translated region, restricted degradation of the polyprotein by protease, and then the release of the structural proteins (Core, E1 and E2) and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B), respectively. Among these structural proteins and non-structural proteins, that is, viral proteins of HCV, Core is a core protein, E1 and E2 are envelope proteins, and non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B) are proteins involved in virus's own replication. NS2 is known to have metalloprotease activity, and NS3 is known to have serine protease activity (at one-third of the N terminal side) and helicase activity (at two-thirds of the C-terminal side). Furthermore, NS4A is a cofactor for protease activity of NS3, and NS5B has been reported to have RNA-dependent RNA polymerase activity. Furthermore, the genome of HCV of genotype 2a has already been reported to have a similar gene structure (see Patent Literature 1).

[0016]

We have constructed RNA capable of autonomous replication using such HCV genome of genotype 2a. Specifically, the HCV-derived replicon RNA of the present invention is an RNA construct, which contains the whole or partial RNA of the HCV genome of genotype 2a and is capable of autonomous replication. [0017]

In this specification, RNA that is prepared by altering the viral genome of HCV and is capable of autonomous replication is referred to as "replicon RNA" or "RNA replicon." RNA that is artificially prepared from HCV of genotype 2a and is capable of autonomous replication is referred to as "replicon RNA derived from HCV of genotype 2a." In this specification, the HCV-derived replicon RNA is also referred to as an HCV-RNA replicon.

F00181

In the present invention, "hepatitis C virus of genotype 2a" or "HCV of genotype 2a" means hepatitis C virus identified as genotype 2a according to the international classification of Simmonds et al. The "hepatitis C virus of genotype 2a" or the "HCV of genotype 2a" of the present invention encompasses not only a virus having naturally occurring HCV genomic RNA, but also a virus having genomic RNA prepared by artificially altering a naturally occurring HCV genomic sequence. Specific examples of HCV of genotype 2a include viruses of JFH-1 strain and the JCH-1 strain (see Patent Literature 1).

T00191

[0020]

Furthermore, "the genomic RNA of hepatitis C virus of genotype 2a" means RNA that comprises the single-stranded (+) strand sense RNA of hepatitis C virus of genotype 2a and has the nucleotide sequence throughout the entire region of its genome. The genomic RNA of hepatitis C virus of genotype 2a is preferably RNA comprising the nucleotide sequence represented by SEQ ID NO: 3 or 5, but is not limited thereto.

In the specification of the present application, "5' untranslated region"

(5'NTR or 5'UTR), "a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein," "a sequence encoding Core protein" (Core region or C region), "a sequence encoding E1 protein" (E1 region), "a sequence encoding E2 protein" (E2 region), "a sequence encoding N2 protein" (NS2 region), "a sequence encoding NS3 protein" (NS3 region), "a sequence encoding NS4A protein" (NS4A region), "a sequence encoding NS4B protein" (NS4B region), "a sequence encoding NS5A protein" (NS5A region), "a sequence encoding NS5B protein" (NS5B region) and "3' untranslated region" (3' NTR or 3' UTR), and other specific regions or sites are determined based on the nucleotide sequence of SEQ ID NO: 3 of the full-length cDNA (JFH-1 clone) encoding the entire region of the genome of the JFH-1 strain, which is HCV of genotype 2a. The nucleotide sequence of SEQ ID NO: 3 can be obtained from the International DNA Data Bank (DDBJ/EMBL/GenBank) by referring to the accession No. AB047639. Specifically, when a particular HCV RNA sequence is aligned with the nucleotide sequence represented by SEQ ID NO: 3, a sequence to be aligned with nucleotides 1 to 340 on the nucleotide sequence represented by SEQ ID NO: 3 is "5' untranslated region" of the RNA, a sequence to be aligned with the nucleotides 3431 to 9442 on the same are a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein, a sequence to be aligned with the nucleotides 3431 to 5323 on the same is "a sequence encoding NS3 protein," a sequence to be aligned with the nucleotides 5324 to 5485 on the same is "a sequence encoding NS4A protein," a sequence to be aligned with the nucleotides 5486 to 6268 on the same is a sequence encoding NS4B protein," a sequence to be aligned with the nucleotides 6269 to 7666 on the same is "a sequence encoding NS5A protein," a sequence to be aligned with the nucleotides 7667 to 9442 on the same is "a sequence encoding NS5B protein," and a sequence to be aligned with the nucleotides 9443 to 9678 on the same is "3' untranslated region." Furthermore, in this case, gaps, additions, deletions, substitutions or the like may be present in the "aligned" sequences. Furthermore, the above "particular HCV" is not limited thereto, and includes the JFH-1 strain or JCH-1 strain, or viral strains that are derivatives thereof.

One embodiment of the HCV RNA-replicon according to the present invention is a replicon RNA comprising a nucleotide sequence containing at least the 5' untranslated region, a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein, and the 3' untranslated region on the genomic RNA of hepatitis C virus of genotype 2a. The replicon RNA may further contain at least one selection marker gene or one reporter gene, and at least one IRES sequence. Furthermore, this replicon RNA may also contain a sequence encoding a viral protein other than NS3, NS4A, NS4B, NS5A and NS5B proteins on the genomic RNA of hepatitis C virus of genotype 2a.

Another preferred embodiment of HCV RNA-replicon according to the present invention is a replicon RNA comprising a nucleotide sequence containing the 5' untranslated region comprising the nucleotide sequence represented by SEQ ID NO: 9 or 10, at least one selection marker gene or reporter gene, the IRES sequence, a sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein on the genomic RNA of hepatitis C virus of genotype 2a, and the 3' untranslated region comprising the nucleotide sequence represented by SEQ ID NO: 11 or 12. In this case the nucleotide sequences represented by SEQ ID NO: 9 and 10 are sequences of the 5' untranslated regions of rSGREP-JFHI (SEQ ID NO: 2), respectively, which are replicon RNAs according to the present invention. Furthermore, the nucleotide sequences represented by SEQ ID NO: 11 and 12 are sequences of the 3' untranslated regions of rSGREP-JFHI (SEQ ID NO: 1) and rSGREP-JCHI (SEQ ID NO: 2), respectively, which are replicon RNAs according to the present invention.

[0023]

A more preferred embodiment of HCV RNA-replicon according to the present invention is a replicon RNA comprised of an RNA comprising the nucleotide sequence represented by SEO ID NO: 1 or 2. Furthermore, a replicon RNA comprising a nucleotide sequence derived from the nucleotide sequence represented by SEO ID NO: 1 or 2 by deletion, substitution or addition of 1 to 50, 1 to 30, 1 to 10, 1 to 6, or 1 to several (2 to 5) nucleotides, and being capable of autonomous replication is also included in the scope of the present invention as a preferred embodiment. In the present invention, "capable of autonomous replication" means that when replicon RNA is introduced into a cell, the replicon RNA allows its own full-length sequence to be replicated within the cell. For example, this ability of autonomous replication can be confirmed by transfecting replicon RNA into Huh7 cells, culturing the Huh7 cells, extracting RNA from the cells in the thus resulting culture product and conducting Northern blot hybridization for the extracted RNA using a probe that can specifically detect the transfected replicon RNA so as to detect the presence of the replicon RNA. However, examples of such a method are not limited thereto. Specific procedures for confirming the ability of autonomous replication can be conducted according to descriptions given in the Examples of this specification such as those for measuring the ability of colony formation, those for confirming the expression of HCV proteins or those for detecting replicon RNA. [0024]

In the present invention, a "selection marker gene" means a gene that can provide a cell with selectivity such that only the cell expressing the gene is selected. A general example of a selection marker gene is an antibiotic resistance gene. In the present invention, preferred examples of a selection marker gene include a neomycin resistance gene, a thymidine kinase gene, a kanamycin resistance gene, a pyrithiamine resistance gene, an adenylyl transferase gene, a Zeocin resistance gene and a puromycin resistance gene. The neomycin

resistance gene is more preferred. However, the selection marker gene in the present invention is not limited to these genes.

[0025]

Furthermore in the present invention, a "reporter gene" means a marker gene encoding a gene product that is a marker for the expression of the gene. General examples of a reporter gene include structural genes of enzymes that catalyze light emitting reaction or color reaction. Preferred examples of the reporter gene in the present invention include a transposon Tn9-derived chloramphenicol acetyltransferase gene, an Escherichia coli-derived β glucuronidase or β galactosidase gene, a luciferase gene, a green fluorescence protein gene, an acquorin gene from jellyfish, and a secreted placental alkaline phosphatase (SEAP) gene. However, the reporter gene in the present invention is not limited to these genes.

[0026]

[0027]

Either only one or both of the above selection marker gene and reporter gene may be contained in replicon RNA.

In the present invention, "IRES sequence" means an internal ribosome entry site that allows translation to be initiated by binding ribosomes within the inside of RNA. Preferred examples of IRES sequence in the present invention include, but are not limited to, EMCV IRES (the internal ribosome entry site of encephalomyocarditis virus), FMDV IRES and HCV IRES. EMCV IRES and HCV IRES are more preferred, and EMCV IRES is the most preferred sequence.

The replicon RNA according to the present invention may further contain a sequence on the genomic RNA of another HCV strain or HCV of another genotype. For example, the replicon RNA may also contain a fragment of HCV genome of genotype 1b. Examples of another HCV strain include, but are not limited to, HCV-1, HCV-H, HC-J1, HCT-18, H77, DK-7, US11, S14, HCT23, HCV-Th, DR1,

DR4, HCT27, S18, SWI, DK9, H90, TD-6EI, S9, HCV-BK, T10, DK1, HC-J4, HCV-J, HK3, HK8, HK5, HCV-G3, IND5, IND8, P10, D1, D3, SW2, T3, S45, SA10, US6, HCV-JK1, HCV-JK4, HCV-JK3, HCV-JK2, HCV-JT, HC-J2, HCV-T, HK4, HC-G9, Z1, Bi, S. I., Cho, J.M., HCV-J6, T4, T9, US10, HC-J5, T2, HC-J7, DK11, SW3, DK8, T8, HC-J8, S83, HK2, HC-J6, HC-J8, BBBE1, HCV-J6, HCV-J8, HD10-2, BR36-9, S52, S54, S2, BR33-1, HK10, DK12, HCV-TR, BA-1, BA-2, DK13, Z1, Z4, Z6, Z7, HK2, SA1, SA4, SA5, SA7, SA13, SA6, NZL1, SA30, EG-13, HCV-K3a/650, ED43, EUH1480, EUHK2, Th580, VN235, VN405, VN004, JK049, JK046, JFH-1, JCH-1, JCH-2, JCH-3, JCH-4, JCH-5, JCH-6, J6CF and H77.

[0029]

The replicon RNA according to the present invention preferably has the 5' untranslated region on the genomic RNA of HCV of genotype 2a on the 5'-most side, and the 3 'untranslated region on the genomic RNA of HCV of genotype 2a on the 3'-most side. A selection marker gene or a reporter gene may be ligated upstream of the IRES sequence, or upstream or downstream of "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein," or inserted in the middle of "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5B protein, NS5B protein, NS5B protein, NS5B protein."

[0030]

The replicon RNA according to the present invention more preferably has the 5' untranslated region on the genomic RNA of HCV of genotype 2a on the 5'-most side, and a selection marker gene or a reporter gene, the IRES sequence and "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5B protein and NS5B protein" downstream of the 5' untranslated region in this order, and the 3' untranslated region on the genomic RNA of HCV of genotype 2a on the 3'-most side.

F00311

Examples of the replicon RNA according to the present invention may

include an RNA containing any foreign gene to be expressed within a cell into which the replicon RNA is introduced, in addition to the sequences as described above. A foreign gene may also be ligated downstream of the 5' untranslated region, or ligated upstream or downstream of a selection marker gene or a reporter gene, or ligated upstream or downstream of "the sequence encoding NS3 protein, NS4A protein, NS5A protein and NS5B protein," or may be inserted in the middle of "the sequence encoding NS3 protein, NS5A protein and NS5B protein, and NS5B protein and NS5B protein and NS5B protein." A replicon RNA containing a foreign gene can express a protein encoded by the foreign gene when it is translated within a cell into which the RNA is introduced. Thus, the replicon RNA containing a foreign gene can be appropriately used also for gene therapy or the like, the purpose of which is to generate a particular gene product within a cell.

The replicon RNA according to the present invention may further contain a ribozyme. A ribozyme is inserted to ligate a selection marker gene, a reporter gene or a foreign gene on the 5' side in the replicon RNA to those located on the 3' side thereof including the IRES sequence and "the sequence encoding NS3 protein, NS4A protein, NS5A protein, NS5A protein and NS5B protein," so that it enables cleavage and separation of the two by the self-cleavage activity of the ribozyme.

In the replicon RNA according to the present invention, the above described selection marker gene, reporter gene, sequences encoding viral proteins on the genomic RNA of hepatitis C virus of genotype 2a, sequences encoding viral proteins of HCV of a genotype other than genotype 2a, a foreign gene or the like are ligated so that they are translated from the replicon RNA in the correct reading frame. Among these sequences, the protein-coding sequences may be ligated to each other via a protease cleavage site and the like, so that after the proteins are expressed as a fusion protein with the polyprotein that is translated from "the sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and

NS5B protein" of hepatitis C virus of genotype 2a, the fusion protein is separated by protease into each protein.

[0034]

2. Preparation of replicon RNA according to the present invention

The HCV RNA-replicon according to the present invention can be prepared using any genetic engineering techniques known by persons skilled in the art. The HCV RNA-replicon can be prepared by, for example, the following method, but the method of preparation is not limited thereto.

[0035]

First, DNA corresponding to the entire region of the genomic RNA of hepatitis C virus of genotype 2a is ligated downstream of an RNA promoter according to a standard procedure so as to prepare a DNA clone. As used herein, "DNA corresponding to RNA" means a DNA having a nucleotide sequence derived from the nucleotide sequence of the RNA by substituting U (uracil) with T (thymine). The above RNA promoter is preferably an RNA promoter contained in a plasmid clone. An example of an RNA promoter is not limited, but T7 RNA promoter is particularly preferred.

[0036]

Next, for the thus prepared DNA clone, for example, the structural region (Core sequence, E1 sequence and E2 sequence) located downstream of the 5' untranslated region and the sequence encoding NS2 protein are substituted with a DNA fragment containing a selection marker gene or a reporter gene and the IRES sequence ligated downstream thereof. In this substitution, portions other than the structural region, such as a fragment on the 3' terminal side of the 5' untranslated region or a part of the sequence encoding NS3 protein may be substituted with a sequence derived from HCV of another genotype.

[0037]

Subsequently, using the DNA clone after the substitution as a template, RNA is synthesized using RNA polymerase. RNA synthesis can be initiated by a

standard procedure from the 5' untranslated region and the IRES sequence. When a template DNA is a plasmid clone, the above DNA region ligated downstream of an RNA promoter is excised by a restriction enzyme from the plasmid clone, and then RNA can be synthesized using the DNA fragment as a template. In addition, preferably the 3' terminus of RNA to be synthesized agrees with the 3' untranslated region of the viral genomic RNA, and no other sequences are added or deleted. The thus synthesized RNA is the replicon RNA according to the present invention. [0038]

3. Preparation of replicon-replicating cells into which replicon RNA from HCV of genotype 2a is introduced

The replicon RNA that is prepared as described above is introduced into cells in which the replicon RNA should be replicated, so that cells wherein the replicon RNA is continuously replicated can be obtained. In this specification, a cell wherein replicon RNA is continuously replicated is referred to as a "replicon-replicating cell."

[0039]

As a cell into which replicon RNA is introduced, any cell can be used, as long as it can be subcultured. Such a cell is preferably a eukaryotic cell, more preferably a human liver-derived cell, a human uterine cervix-derived cell or a human fetal kidney-derived cell, and further preferably any cell selected from the group consisting of Huh7 cells, HepG2 cells, IMY-N9 cells, HeLa cells and 293 cells. As these cells, commercially available cells may be utilized, these cells may be obtained from cell depositories, or cell lines established from any cells (e.g., cancer cells or stem cells) may also be used.

As the above cells, cells that can be mass-cultured are preferably used for the purpose of the mass production of HCV proteins, such as in the case of vaccine production. From such a viewpoint, the cells are preferably those other than Huh7 cells.

[0041]

Introduction of replicon RNA into cells can be performed using any technique known by persons skilled in the art. Examples of such an introduction method include electroporation, a particle gun method, a lipofection method, a calcium phosphate method, a microinjection method and a DEAE sepharose method. The method using electroporation is particularly preferred.

A replicon RNA of interest may be introduced alone, or may be introduced after it is mixed with other nucleic acids. To vary the quantity of replicon RNA while keeping RNA quantity to be introduced at a certain level, the replicon RNA of interest is mixed with total cellular RNA extracted from cells into which the RNA is introduced, and then the mixture is used for introduction into cells. The quantity of replicon RNA to be used for introduction into cells may be determined depending on the introduction method employed, and is preferably between 1 picogram and 100 micrograms, and more preferably between 10 picograms and 10 micrograms.

[0043]

When replicon RNA containing a selection marker gene or a reporter gene is used for introduction into cells, cells wherein the replicon RNA is introduced and continuously replicated can be selected utilizing the expression of the selection marker gene or the reporter gene. Specifically, for example, such cells into which replicon RNA has been introduced may be cultured in media whereby the cells can be selected by the expression of the selection marker gene or the reporter gene. As an example, when replicon RNA contains a neomycin resistance gene as a selection marker gene, cells into which replicon RNA has been intracellularly introduced are seeded into a culture dish. After 16 to 24 hours of culture, G418 (neomycin) is added to the culture dish at a concentration of 0.05 milligrams/milliliter to 3.0 milligrams/milliliter. The cells are continuously cultured for preferably 10 days to 40 days and more preferably 14

days to 28 days after seeding, while exchanging the culture solution twice a week. Next, surviving cells are stained with crystal violet, so that cells into which the replicon RNA has been introduced and is being continuously replicated can be selected as formed colonies.

f00441

Cloned cells can be obtained from the formed colonies by cloning surviving cells by a standard procedure, and then continuing the culture of the cells. The thus obtained cell clone wherein the replicon RNA of interest is continuously replicated is referred to as "a replicon-replicating cell clone" in this specification.

[0045]

Regarding the established cell clone, detection of a replicon RNA that has been replicated from the introduced replicon RNA in the cell clone, confirmation of the presence or the absence of the incorporation of a selection marker gene or a reporter gene in the introduced replicon RNA into a host genomic DNA, and confirmation of the expression of an HCV protein are preferably carried out to confirm the fact that a replicon RNA of interest is actually and continuously replicated.

[0046]

A replicon RNA that has been replicated from the introduced replicon RNA in the cell clone (in this specification, hereinafter conveniently referred to as "replicated replicon RNA") may be detected according to any RNA detection method known by persons skilled in the art. For example, detection can be performed by conducting the Northern hybridization method for total RNA extracted from the cell clone using as a probe a DNA fragment specific to the introduced replicon RNA.

F00471

Furthermore, the presence or the absence of the incorporation of a selection marker gene or a reporter gene in the introduced replicon RNA into a

host genomic DNA can be confirmed by, for example, performing PCR for the host genomic DNA extracted from the cell clone to amplify at least a part of the selection marker gene or the reporter gene, and then confirming the presence or the absence of the amplified product. However, examples of relevant methods are not limited thereto. A cell clone for which the amplified product is confirmed is considered to have a selection marker gene or a reporter gene incorporated in the host genome. Thus, regarding the cell clone, the replicon RNA itself may not be continuously replicated within the cell. In this case, whether or not the replicon RNA is continuously replicated can be confirmed by conducting an experiment to confirm the expression of an HCV protein, as described below.

[0048]

The expression of an HCV protein can be confirmed by, for example, causing an antibody against an HCV protein to be expressed from the introduced replicon RNA and to react with a protein extracted from a cell clone. This method can be conducted by any protein detection method known by persons skilled in the art. Specifically, for example, a protein sample extracted from the cell clone is blotted onto a nitrocellulose membrane, with which an anti-HCV protein antibody (e.g., an anti-NS3-specific antibody or an antiserum collected from a hepatitis C patient) is reacted, and then the anti-HCV protein antibody is detected. If the HCV protein is detected among proteins extracted from the cell clone, it can be concluded that this cell clone continuously replicate HCV-derived replicon RNA to express the HCV protein.

[0049]

As described above, cell clones confirmed to continuously replicate a replicon RNA of interest (replicon-replicating cell clones) can be obtained. Furthermore in the present invention, replicon RNA can be obtained by any method known by persons skilled in the art, for example, by extracting RNA from the replicon-replicating cell, and then separating replicon RNA from the RNA by an electrophoresis method. The present invention also relates to such a method of producing replicon RNA. Moreover, preferably, the replicon-replicating cell according to the present invention can be used for producing HCV proteins. Persons skilled in the art can obtain HCV proteins from the replicon-replicating cells according to any standard method. Specifically, for example, a viral protein of hepatitis C virus of genotype 2a can be produced by culturing replicon-replicating cells, collecting proteins from the resulting culture product (including cultured cells and culture media) by a standard procedure, and then selectively obtaining viral proteins from the proteins by detection or the like using an anti-HCV protein antibody.

[0050]

Moreover, when the replicon-replicating cell according to the present invention continuously replicates replicon RNA containing a foreign gene, a protein encoded by the foreign gene can be obtained by the expression thereof using the replicon-replicating cell. Specifically, for example, the protein encoded by a foreign gene can be obtained by culturing replicon-replicating cells, collecting proteins from the resulting culture product (including cultured cells and culture media) by a standard procedure, and then selectively obtaining the protein from among the proteins by detection or the like using an antibody against the protein of interest.

[0051]

 Introduction of mutation that increases replication efficiency into replicon RNA from HCV of genotype 2a

Mutation producing enhancement of replication efficiency frequently takes place in the replicon RNA that is replicated or generated in the repliconreplicating cell (replicated replicon RNA) according to the present invention.

Such a mutation may be an adaptive mutation.

Utilizing this fact, introduction of a mutation enhancing replication efficiency into the replicon RNA according to the present invention can be promoted in the present invention.

[0052]

Specifically, the step comprising obtaining a first replicated replicon RNA by extraction or the like from a first replicon-replicating cell (preferably, a replicon-replicating cell, wherein the replicon RNA according to the present invention has been introduced), and then re-introducing the first replicated replicon RNA into another cell to prepare a second replicon-replicating cell is performed repeatedly once or more, preferably 1 to 10 times, more preferably 1 to 5 times, and further preferably 1 to 2 times, so that the mutation increasing replication efficiency can be introduced at a high frequency into the replicon RNA within the replicon-replicating cells.

[0053]

As a cell into which a replicated replicon RNA is re-introduced, any cell can be used. Such a cell is preferably derived from a biological species that is the same as that of a cell wherein replicon RNA is introduced at the beginning, more preferably derived from the same tissue derived from the same biological species as that of a cell wherein replicon RNA is introduced at the beginning, and further preferably of a cell line that is the same as that for a cell wherein replicon RNA is introduced at the beginning.

[0054]

Therefore in the present invention, using the above method, replicon RNA wherein the mutation increasing replication efficiency is introduced can be produced. Specifically, the step comprising obtaining a first replicated replicon RNA by extraction or the like from a first replicon-replicating cell (preferably, a replicon-replicating cell, into which the replicon RNA according to the present invention has been introduced), and then re-introducing the first replicated replicon RNA into another cell so as to prepare a second replicon-replicating cell is performed repeatedly once or more, preferably 1 to 10 times, more preferably 1 to 5 times, and further preferably 1 to 2 times. Subsequently, the replicated replicon RNA is obtained by extraction or the like from the replicon-replicating

cell finally obtained at the end of the repeated steps, so that replicon RNA with increased replication efficiency can be produced.

[0055]

In the present invention, the replication efficiency of a replicon RNA can be increased at least 2 times, preferably 10 to 100 times, and more preferably 100 to 10000 times by the above method.

[0056]

Regarding the replicon RNA that is produced by such a method so as to have increased replication efficiency, the nucleotide sequence is preferably determined by a known method, for example, by obtaining cDNA by reverse transcription PCR and subjecting such cDNA to sequencing. Furthermore, the thus determined nucleotide sequence or the amino acid sequence encoded by the nucleotide sequence is compared with the nucleotide sequence of replicon RNA that had been introduced at the beginning into cells, so that adaptive mutation can be identified. As adaptive mutation increasing replication efficiency, in particular, nonsynonymous substitution that mutates an amino acid in a viral protein encoded by replicon RNA is preferred.

[0057]

The present invention also provides a method whereby the replicon RNA of hepatitis C virus of genotype 2a having increased replication efficiency can be produced by introducing the thus identified adaptive mutation into replicon RNA, the replication efficiency of which is to be increased, by a standard procedure.

The replicon RNA that is produced as described above so as to have increased replication efficiency can be used for producing replicon RNA in large quantity within cells that have been used for the method.

[0059]

The replication efficiency of the replicon RNA according to the present invention can be determined by a method known by persons skilled in the art. For example, it can be determined according to the following method. Replicon RNAs are transfected in quantities of 0.0001, 0.0003, 0.01, 0.003, 0.01, 0.03, 0.1, 0.3 and 1.0 micrograms, respectively, into Huh7 cells, selective culture with G418 is performed for 21 days in a method similar to the above experimental techniques, and then the number of colonies formed (number of colonies) is counted. The quantity of RNA introduced is compared with the number of colonies formed to determine the range of the quantity of the replicon RNA introduced, within which colony formation increases in a quantity-dependent manner. The number of colonies formed within the range is divided by the quantity of RNA introduced, and the resulting value is regarded as the colony forming activity per microgram. This equation is as follows.

Colony forming activity [(Colony Forming Unit, or CFU)/microgram] =
Number of colonies formed [colony] / quantity of RNA introduced [microgram]
[0060]

The thus calculated colony forming activity is regarded as a value representing the replication efficiency of replicon RNA introduced. Specifically, the higher the colony forming activity, the higher the replication efficiency of the replicon RNA. In addition, the replication efficiency of replicon RNA can also be shown via a colony-forming ability that is represented by the number of copies of the replicon RNA introduced per formed colony. That is, in this case, the ability can be calculated according to the following equation.

Colony forming ability = number of copies of replicon RNA introduced [copy] / number of formed colonies [colony] [0061]

5. Other embodiments of the present invention

The replicon RNA-replicating cell according to the present invention can also be used as a test system for, for example, screening for a substance that promotes or suppresses the replication of hepatitis C virus. Specifically, for example, replicon replicating cells are cultured in the presence of a test substance,

replication of the replicon RNA in the resulting culture product is detected, and then whether or not the test substance promotes or suppresses the replication of the replicon RNA is determined, so that a substance that promotes or suppresses the replication of hepatitis C virus can be screened for. In this case, detection of the replication of the replicon RNA in the resulting culture product may be conducted by detecting the quantity of, or the presence or the absence of, the replicon RNA in the RNAs extracted from the replicon RNA-replicating cell, or by detecting the quantity of, or the presence or the absence of, HCV protein contained in the proteins in the culture product or in the replicon RNA-replicating cells contained in the culture product.

100621

[0063]

Such a test cell system using the replicon RNA-replicating cells according to the present invention may be aimed at producing or evaluating a therapeutic agent or a diagnostic agent for treating hepatitis C virus infection. Specific examples of such purposes include the following examples.

(1) Search for a substance suppressing the proliferation of HCV of genotype 2a

Examples of a substance suppressing the proliferation of HCV of genotype 2a include organic chemicals directly or indirectly affecting the proliferation of HCV of genotype 2a, and antisense oligonucleotides directly or indirectly affecting the proliferation of HCV or the translation of HCV proteins by hybridizing to a target sequence in the HCV genome of genotype 2a or a complementary strand thereof.

(2) Evaluation of various substances having antiviral action in cell culture

Examples of the various substances include substances obtained through rational drug design or high throughput screening (e.g., an isolated and purified enzyme) and the like.

(3) Identification of a new target for attack for treating patients infected with HCV of genotype 2a

To identify a host cellular protein that plays an important role in proliferation of HCV virus, for example, the replicon-replicating cell according to the present invention can be used.

- (4) Evaluation of the ability of HCV virus to acquire resistance against a drug or the like and identification of mutation concerning such resistance
- (5) Production of a viral protein as an antigen that can be used for developing, producing and evaluating a diagnostic agent or a therapeutic agent for hepatitis C virus infection
- (6) Viral genome replication system for producing HCV virus or virus-like particles that can be used for developing, producing and evaluating a diagnostic agent or a therapeutic agent for hepatitis C virus infection
- (7) Production of a vaccine antigen that can be used as a vaccine against HCV of genotype 2a
- (8) Production of hepatic cell-directed genetic vector that is used after the incorporation of a foreign gene therein for gene therapy

[Examples]

[0064]

The present invention will be described more specifically based on the following examples and drawings. However, the technical scope of the present invention is not limited by these examples.

[0065]

[Example 1] Preparation of replicon RNA

(A) Construction of expression vector

DNA corresponding to the entire region of viral genome of hepatitis C virus JFH-1 strain (genotype 2a) that had been separated from patients with fullminant hepatic failure was obtained from a JFH-1 clone containing the full-length genomic cDNA of the virus strain. The DNA was inserted downstream of T7 RNA promoter sequence that had been inserted in pUC19 plasmid. The thus constructed plasmid DNA is hereinafter referred to as pJFH1. Similarly, DNA

corresponding to the entire region of viral genome of hepatitis C virus JCH-1 strain (genotype 2a) that had been separated from patients with chronic hepatitis was obtained from a JCH-1 clone containing the full-length genomic cDNA of the virus strain. The DNA was inserted downstream of the T7 RNA promoter sequence that had been inserted in pUC19 plasmid. The thus constructed plasmid DNA is hereinafter referred to as pJCH1. In addition, the preparation of the above JFH1 clone and JCH-1 clone is described in Patent Literature 1 and Non Patent Literature 3. Moreover, the nucleotide sequence of the full-length cDNA of JFH-1 clone was registered at the International DNA Data Bank (DDBJ/EMBL/GenBank) under accession No. AB047649, and the nucleotide sequence of the full-length cDNA of the JCH-1 clone under accession No. AB047640.

[0066]

The structures of the thus constructed plasmid DNA pJFH1 and pJCH1 are shown in the upper section of Fig. 1. "T7" represents T7 RNA promoter, and "G" represents dGTP inserted upstream of the 5' end of the inserted JPH-1- or JCH-1-derived DNA and downstream of the 3' end of T7 RNA promoter sequence. A region from "5' NTR" to "3' NTR" is DNA corresponding to the entire genomic region of hepatitis C virus.

[0067]

Next, the structural regions and a part of the non-structural regions of plasmid DNA pJFH1 and pJCH1 were substituted with a neomycin resistance gene (neo; also referred to as a neomycin phosphotransferase gene) and EMCV-IRES (internal ribosome entry site of encephalomyocarditis virus), thereby constructing plasmid DNA pSGREP-JFH1 and pSGREP-JCH1, respectively (lower section of Fig. 1). This construction procedure was conducted according to a previous report (Non Patent Literature 7). Specifically, plasmid pJFH1 and pJCH1 were cleaved with restriction enzymes Age I and Cla I, and between the Age I and Cla I restriction sites, the following fragments were inserted to be ligated; a fragment

was prepared by binding of a sequence ranging from 5' NTR to Core region derived from pJFH-1 with the neomycin resistance gene derived from pRSV5NEO by PCR amplification and then cleaving it with restriction enzymes Age I and Pme I, and, a fragment was prepared by binding of sequences ranging from EMCV IRES to NS3 region by PCR amplification and then cleaving it with restriction enzymes Pme I and Cla I.

[8800]

Moreover, a mutation that mutates an amino acid motif GDD to GND, corresponding to the active center of RNA polymerase encoded by the NS5B region, was introduced into the NS5B region in pSGREP-JFH1, thereby preparing a mutant plasmid clone pSGREP-JFH1/GND.

[0069]

Moreover, a mutation that results in the deletion of a sequence of 10 continuous amino acids containing an amino acid motif GDD corresponding to the active center of RNA polymerase encoded by the NS5B region was introduced into the NS5B region in pSGREP-JFH1, thereby preparing a mutant plasmid clone pSGREP-JFH1/dGDD.

[0070]

The above-prepared mutant clones pSGREP-JFH1/GND and pSGREP-JFH1/dGDD cannot express active NS5B protein, which is required for the replication of replicon RNA, because the amino acid sequence of the active site of NS5B protein encoded by these clones has mutated.

[0071]

(B) Preparation of replicon RNA

To prepare template DNA for use in synthesis of replicon RNA, the aboveconstructed expression vectors pSGREP-JFH1, pSGREP-JCH1, pSGREP-JFH1/GND and pSGREP-JFH1/dGDD were each cleaved with a restriction enzyme Xba I.

100721

Subsequently, 10 to 20 µg each of these Xba I-cleaved fragments was contained in 50 µl of a reaction solution, and then further treated by 30 minutes of incubation at 30°C with 20 U of Mung Bean Nuclease. Mung Bean Nuclease is an enzyme catalyzing a reaction for selectively degrading a single-stranded portion of double-stranded DNA. Generally, when RNA synthesis is performed using directly the above Xba I-cleaved fragment as a template, a replicon RNA having four nucleotides of CUGA, a part of the recognition sequence of Xba I, excessively added to the 3' terminus would be synthesized. Hence, in this example, Xba I-cleaved fragments were treated with Mung Bean Nuclease, so as to remove the four nucleotides of CUGA from the fragments. The solutions containing Xba I-cleaved fragments, which had been treated with Mung Bean Nuclease, were treated to remove proteins according to a general method, so that Xba I-cleaved fragments, from which the four nucleotides of CUGA had been removed, were purified and used as template DNAs.

Γ00731

Next, from the template DNA, RNA was synthesized in vitro using T7 RNA polymerase. For this RNA synthesis, MEGAscript from Ambion, Inc. was used. Reaction was carried out using 20 μ l of a reaction solution containing 0.5 to 1.0 micrograms of the template DNA according to the instructions of the manufacturer.

[0074]

After completion of RNA synthesis, DNase (2 U) was added to the reaction solution to conduct reaction at 37°C for 15 minutes. RNA extraction using acidic phenol was further performed to remove the template DNA. RNAs (replicon RNAs) synthesized in this manner from the above template DNAs derived from pSGREP-JFH1, pSGREP-JCH1, pSGREP-JFH1/GND and pSGREP-JFH1/dGDD were respectively named rSGREP-JFH1, rSGREP-JCH1, rSGREP-JFH1/GND and rSGREP-JFH1/GDD. Regarding the nucleotide sequences of these replicon RNAs, the nucleotide sequence of rSGREP-JFH1 is shown in SEQ ID NO: 1 and

Fig. 2, that of rSGREP-JCH1 is shown in SEQ ID NO: 2 and Fig. 3, that of rSGREP-JFH1/GND is shown in SEQ ID NO: 7, and that of rSGREP-JFH1/dGDD is shown in SEQ ID NO: 8.

[0075]

[Example 2] Establishment of replicon-replicating cell clone

(C) Transfection of replicon RNA, determination of colony-forming ability of transfected cells and establishment of cell clones

Each of the above-synthesized replicon RNAs (rSGREP-JFH1, rSGREP-JCH1, rSGREP-JFH1/GND) and rSGREP-JFH1/dGDD) was mixed in different quantities with total cellular RNA extracted from Huh7 cells so as to have a total RNA quantity of 10 μg. Subsequently, the mixed RNA was introduced into Huh7 cells by the electroporation method. The Huh7 cells subjected to the electroporation treatment were seeded into culture dishes, and then cultured for 16 hours to 24 hours. G418 (neomycin) was then added to the culture dishes at different concentrations. Thereafter, culture was continued while exchanging the culture solutions twice a week. After 21 days of culture following seeding, surviving cells were stained with crystal violet. The number of stained colonies was counted, and then the number of colonies obtained per μg of the transfected replicon RNA was calculated.

[0076]

For rSGREP-JFH1 or rSGREP-JCH1-transfected cells, for which colony formation had been observed, colonies of the surviving cells were further cloned from the above culture dishes after 21 days of culture, and were continuously cultured. By such cloning of colonies, several strains of cell clones could be established.

[0077]

For the established cell clones, detection of the replicated replicon RNA, confirmation of the presence or the absence of the incorporation of the neomycin resistance gene into the host genomic DNA, and confirmation of the expression of

HCV proteins were performed as described later, in Example 4. Cell clones for which the replication of the replicon had been confirmed in the cells were regarded as replicon-replicating cell clones.

[0078]

(D) Colony-forming ability in each transfected cell

As a result of the above transfection, for rSGREP-JFH1-transfected Huh7 cells, the colony-forming ability per µg of the transfected replicon RNA was 94700 CFU (Colony Forming Unit)/µg-RNA when G418 concentration was 1.0 mg/ml (the left column in Fig. 4). In contrast, colony formation was not observed in the Huh7 cells, into which rSGREP-JFH1/dGDD and rSGREP-JFH1/GND had each been transfected (the central column and the right column in Fig. 4). This suggests that the colony-forming ability confirmed for the Huh7 cells, into which rSGREP-JFH1 replicon RNA had been transfected, depends on the activity of NS5B (RNA polymerase) expressed by rSGREP-JFH1. Specifically, it was considered that in cells that had formed colonies, rSGREP-JFH1 replicon RNA autonomously replicated due to the action of NS5B expressed by rSGREP-JFH1, and the neomycin resistance gene was continuously expressed to maintain G418 resistance, so that cell growth was enabled.

On the other hand, in the Huh7 cells, into which rSGREP-JCH1 had been transfected, no colony formation was observed in the case of 1 to 0.5 mg/ml G418 concentrations (Fig. 5). When G418 concentration was lowered to 0.25 mg/ml, colony formation was observed in the Huh7 cells, into which rSGREP-JCH1 had been transfected as well.

[0080]

Furthermore, Xba I-cleaved fragment of the expression vector pSGREP-JFH1 obtained in (B) above was used as a template DNA for RNA synthesis without treating the fragment with Mung Bean Nuclease, so as to synthesize replicon RNA. This replicon RNA was transfected to Huh7 cells in a manner similar to that in (C) above. The replicon RNA that had been prepared without performing Mung Bean Nuclease treatment had the four nucleotides of CUGA excessively added to the 3' terminus.

[0081]

As a result, the colony-forming ability of the Huh7 cells, into which the replicon RNA prepared without treatment with Mung Bean Nuclease had been transfected, decreased to 512 CFU/µg·RNA (the left side in Fig. 6). This result revealed that the sequence on the 3' terminus of the replicon RNA affects the colony-forming ability of the transfected cells.

[0082]

[Example 3]

(E) Re-transfection of replicated replicon RNA derived from replicon-replicating cells

From the replicon-replicating cell clones that had been established by transfection of rSGREP-JFH1 into Huh7 cells according to descriptions of Example 2, total RNA was extracted by a standard procedure. The number of copies of the replicated replicon RNA contained in the cellular RNA was determined by Northern blot analysis and a quantitative RT-PCR method.

[0083]

Northern blot analysis was performed according to the description in Molecular Cloning, A laboratory Manual, 2nd edition, J. Sambrook, E. F. Fritsch, T. Maniatis, Cold Spring Harbor Laboratory Press (1989). Specifically, RNA extracted from the cells was subjected to denaturing agarose electrophoresis. After electrophoresis, the RNA was transferred onto a positively charged nylon membrane. The ³²P-labeled DNA or RNA probe prepared from pSGREP-JFH1 was hybridized to the RNA transferred to the membrane as described above. Next the membrane was washed, and then exposed to a film, so as to detect a replicon-specific RNA band.

100841

Detection of the replicon RNA by quantitative RT-PCR was conducted by detecting the 5' untranslated region RNA within HCV RNA according to Takeuchi T, Katsume A, Tanaka T, Abe A, Inoue K, Tsukiyama-Kohara K, Kawaguchi R, Tanaka S and Kohara M., Real-time detection system for quantification of Hepatitis C virus genome, Gastroenterology 116: 636-642 (1999). Specifically, the replicon RNA contained in RNA extracted from the cells was amplified by PCR using synthetic primers: R6-130-S17, 5'-CGGGAGAGCCATAGTGG-3' (SEQ ID NO: 13) and R6-290-R19, 5'-AGTACCACAAGGCCTTTCG-3' (SEQ ID NO: 14); TaqMan Probe; R6-148-S21FT, 5'-CTGCGGAACCGGTGAGTACAC-3' (SEQ ID NO: 15) and an EZ rTth RNA PCR kit, and then detected using an ABI Prism 7700 sequence detector system.

[0085]

Next, aliquots of total cellular RNAs extracted from clone 6 (among the above-mentioned replicon-replicating cell clones) and pool clones (prepared by collecting replicon-replicating cells that had formed colonies from whole one dish and culturing them) were each introduced into another Huh7 cells by retransfection. Total cellular RNA used for the transfection was prepared to contain 1x107 copies of replicon RNA based on the number of copies of the above-determined replicon RNA. Transfection was performed as described in (C) above, and then selective culture was performed under G418 concentration conditions of 1 mg/ml. Thus, the colony formation of the replicon-replicating cells was observed (Fig. 7). The colony-forming ability in this case was 1 colony or more per 1x106 copies of the replicon RNA used for transfection, when it was calculated from the number of colonies obtained.

[0086]

On the other hand, the number of copies of in vitro synthetic RNA that had been synthesized in vitro using pSGREP-JFH1 as a template and T7 RNA polymerase was approximately 2x10¹¹ copies/µg-RNA, when calculated based on the weight and the length of the RNA. The colony-forming ability in the case of

using the in vitro synthetic RNA for transfection in a manner similar to the above method was 1 colony per 5×10^{7} copies. These results revealed that when RNA derived from cells extracted from replicon-replicating cells and in vitro synthetic RNA were each transfected to Huh7 cells as replicon RNA in the same number of copies, the use of the replicon RNA replicated within Huh7 cells resulted in colony-forming ability approximately 50 times higher than that of the in vitro synthetic RNA.

[0087]

[Example 4]

(F) Detection of replicon RNA

According to (E) above, cell clones [clones Nos. I to 11] were established by retransfection of total RNA that had been obtained from the repliconreplicating cell clone established by transfection of rSGREP-JFH1 to Huh7 cells to another Huh7 cells. From the established cell clones and pool clones (prepared by collecting cell clones that had formed colonies from whole one dish and then culturing them), respectively, total RNAs were extracted by an acidic phenol extraction method. Subsequently the total RNAs were analyzed by the Northern blot method using a pSGREP-JFH1-specific probe as a probe. As control, total RNA extracted similarly from untransfected Huh7 cells (in Fig. 8, denoted as "Huh7"), a sample prepared by adding 10⁷ copies of replicon RNA synthesized in vitro to the total RNA extracted from Huh7 cells (in Fig. 8, denoted as "10⁸") prepared by adding 10⁸ copies of replicon RNA synthesized in vitro to the total RNA extracted from Huh7 cells, were used. In Fig. 8, 1 to 11 represent cell clone Numbers.

[8800]

As a result, RNA of approximately the same size as that of rSGREP-JFH1 was detected using a pSGREP-JFH1-specific probe (Fig. 8). Thus, it was confirmed that the replicon RNA from rSGREP-JFH1 that had been transfected at the beginning replicated and proliferated within the cell clones. In addition, it

was shown that the cell clones differed from each other in the quantity of the replicated replicon RNA. In Fig. 8, for example, clones 2, 6, 9 and 10 contained high quantities of the replicated replicon RNA, and clones 4, 8 and 11 contained low quantities of the replicated replicon RNA.

[0089]

(G) Confirmation of the presence or the absence of the incorporation of a neomycin resistance gene into genomic DNA

For the cell clones that had been obtained by retransfection of replicon RNA as described in Example 3, PCR amplification was performed using neomycin resistance gene-specific primers; sense primer, NEO-S3: 5'-AACAAGATGGATTGCACGCA-3' (SEQ ID NO: 16) and antisense primer, NEO-R: 5'-CGTCAAGAAGGCGATAGAAG-3' (SEQ ID NO: 17), and the host cellular genomic DNA extracted from each of the cell clones as a template, in order to confirm that the resistance of each of the cell clones against G418 was not due to the incorporation of the neomycin resistance gene into the genome. The cell clones used herein were the cell clones Nos. 1 to 8 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA (rSGREP-JFH1-derived cell clones Nos. 1 to 8), and cell clones Nos. 1 to 6 obtained by retransfection of rSGREP-JCH1-derived replicated replicon RNA (rSGREP-JCH1-derived cell clones Nos. 1 to 6). As a result, as shown in Fig. 9, in the eight examined rSGREP-JFH1-derived cell clones, positive clones showing the amplification of the neomycin resistance gene were not observed. For rSGREP-JCH1-derived cell clones, only 1 out of the 6 examined clones was positive (in Fig. 9, lane 3 in the right photograph). It was considered that this positive clone had acquired G418 resistance by the incorporation of the neomycin resistance gene in rSGREP-JCH1derived replicated replicon RNA into the genomic DNA of the host cells. Thus, in the positive clone, unlike other clones, it was thought that the replicon RNA itself did not autonomously replicate within the cells. This was confirmed by the results of the experiment shown in the next (H) that no HCV proteins were detected from the positive clone.

[0090]

(H) Detection of HCV protein

Protein was extracted from rSGREP-JFH1- and rSGREP-JCH1-transfected cell clones by a standard procedure, and then analyzed by SDS-PAGE and Western blot method (Fig. 10). The examined cell clones were the same as those used in (G) above: rSGREP-JFH1-derived cell clones Nos. 1 to 8 and rSGREP-JCH1derived cell clones Nos. 1 to 6. In addition, a cellular extract from the cell obtained by transiently transfecting expression plasmid DNA containing NS3 gene into Huh7 cells was regarded as a positive control (NS3 protein). Furthermore, a protein extracted from the untransfected Huh7 cells was used as a negative control. A protein sample extracted from each cell clone was blotted onto a PVDF membrane (Immobilon-P, Millipore), and then detection of NS3 protein encoded by replicated replicon RNA was performed using anti-NS3-specific antibody (provided by Dr. Moradpour; Wolk B, et al, J. Virology. 2000, 74: 2293-2304). As shown in Fig. 10, in rSGREP-JFH1-derived cell clones Nos. 1 to 8 and rSGREP-JCH1-derived cell clones Nos. 1, 2 and 4 to 6, proteins of the same size as those of the positive control were detected. In rSGREP-JCH1-derived cell clone No. 3 (the clone detected as a positive clone in (G) above), no expression of NS3 protein was detected. That is, in rSGREP-JCH1-derived cell clone No. 3, no replication of replicon RNA was confirmed. NS3 protein was not detected in the untransfected Huh7 cells, revealing that in cell clones wherein NS3 protein was detected, the transfected replicon RNA autonomously replicated so that NS3 protein was expressed.

[0091]

Moreover, by the use of the serum of a hepatitis C patient as an antibody, the expression of NS5a protein from the replicon RNA was also confirmed in each cell clone for which the expression of NS3 protein had been confirmed as described above.

[0092]

Based on the results of (G) and (H) above, it was confirmed that replicon RNAs were replicated in the cell clones established by transfection of the replicon RNA.

[0093]

[Example 5]

(I) Analysis of adaptive mutation

According to descriptions of Example 3, total RNA obtained from the replicon-replicating cell clones established through the transfection of rSGREP-JFH1 into Huh7 cells was re-transfected to another Huh7 cells, thereby establishing 21 cell clones. Total RNA was extracted from each of these cell clones by a standard procedure. cDNA corresponding to the replicon RNA was synthesized using the total RNA as a template, reverse transcriptase Superscript II (Invitrogen) and primer 9641R-IH (5'-GCACTCTTGCAGTCATGCGGCTCACGGAC-3' (SEQ ID NO: 18)). The composition of a reaction solution for the synthesis of cDNA by reverse transcription reaction is as shown below.

[0094]

Composition of Reaction Solution	Fluid Volume (µl)
5x 1st strand Buffer	4
2 mM dNTP	5
0.1 M DTT	1
9651R-IH primer (100 μM)	1
DW (distilled water)	6.5
Sample RNA (2 mg/mL)	1
RNasin (Promega) (40 U/μL)	0.5
Superscript II RT (Invitrogen)	L
Total	20 μ1
[0095]	

In cDNA synthesis reaction, the above reagents other than RNasin and Superscript II were mixed to prepare a first reaction solution. The solution was heated at 90°C for 3 minutes, and then cooled on ice. Subsequently, RNasin and Superscript II were added to this reaction solution, and then the solution was allowed to react at 42°C for 1 hour, followed by another reaction at 70°C for 15 minutes.

[0096]

Furthermore, PCR amplification was performed using the thus obtained cDNA together with five primer sets by the following procedures, so that DNA amplification fragments covering almost all the regions of the replicon RNA were obtained. The primer sets used and regions amplified by each primer set are shown in Table 1 and Table 2 below.

[0097] [Table 1]

Designation of	Primer set		Amplified region
amplified fragment	Primer 1	Primer 2	
Α/	42S-IH	433R-neo	41 - 470
B/	C/S17ssp	4680R-IH	28 - 3026
C/	4534S-IH	7279R-IH	2880 - 5625
D/	7198S-IH	9367R-IH	5544 - 7713
E/	9247S-NF	9576R-NF	7597 - 7960

In Table 1, an amplified region is represented by nucleotide numbers in rSGREP-JFH1 (SEQ ID NO: 1) that the region corresponds to.

[Table 2]

Primer	Nucleotide sequence (5'→3')	SEQ ID NO:
designation		
42S-IH	CCCCTGTGAGGAACTACTGTCTTCACGC	SEQ ID NO: 19
C/S17ssp	CCGGGAGAGCCATAGTGGTCTGCG	SEQ ID NO: 20
4534S-IH	CCACTCAAAGAAAAAGTGTGACGAGCTCGC	SEQ ID NO: 21
7198S-IH	GGCTTGGGCACGGCCTGA	SEQ ID NO: 22
9247S-NF	GCGGTGAAGACCAAGCTCAAACTCACTCCA	SEQ ID NO: 23

433R-neo	LAGAACCTGCGTGCAATCCATC	SEQ ID NO: 24
4680R-1H	CCCGTCATGAGGGCGTCGGTGGC	SEQ ID NO: 25
	ACCAGCAACGGTGGGCGGTTGGTAATC	SEQ ID NO: 26
7279R-IH	GGCACGCGACACGCTGTG	SEQ ID NO: 27
9367R-RI	AGCTAGCCGTGACTAGGGCTAAGATGGAGC	SEQ ID NO: 28
9576R-NF	AGCIAGCCGIGACIAGGGCIAIG	

[0098]

The composition of a reaction solution in this PCR reaction is as follows.

Composition of Reaction Solution	Fluid Volume (µl)
Primer 1 (10 μM)	1.0
Primer 2 (10 µM)	1.0
2.5 mM dNTPs	5.0
10x LA Buffer	5.0
MgCl ₂ (25 mM)	5.0
LA Taq (TAKARA) (5 U/μl)	0.3
DW (distilled water)	30.7
Template cDNA	2.0
Total	50 μΙ

[0099]

In addition, PCR reaction conditions are as follows: 95°C for 2 minutes; 35 cycles of 98°C for 10 seconds and then 68°C for 8 minutes; and 72°C for 7 minutes; after which the temperature is kept at 4°C.

[0100]

The nucleotide sequence of each PCR product obtained as described above was determined, and then the RNA sequence corresponding to the DNA sequence was compared with the sequence of rSGREP-JFH1. The results are shown in Table 3.

[0101]

[Table 3]

Region	Synonymous substitution	Nonsynonymous substitution	Total number of mutations
NS3	0	5	5
NS4A	0	2	2
NS4B	0	3	3
NS5A	0	7	7
NS5B	3	5	8
Total	3	22	25

[0102]

As shown in Table 3, total number of nucleotide mutations observed in 21 cell clones was 25. 22 of these mutations were nonsynonymous substitutions inducing amino acid mutation. Types of these mutations are as shown in Table 4. In addition, the positions of these mutations in the non-structural region are shown in Fig. 11.

[0103] [Table 4]

Clone	Mutation site			
designation	Nucleotide No.	Nucleotide mutation	Amino acid mutation	Amino acid No.
C1	7098	$A \Rightarrow G$	None	
	7157	$A \Rightarrow G$	$Y \Rightarrow C$	2824
C2	4955	$c \Rightarrow u$	$A \Rightarrow V$	2090
C3	4936	$A \Rightarrow G$	$T \Rightarrow A$	2084
	5000	$A \Rightarrow G$	$Y \Rightarrow C$	2105
	7287	$A \Rightarrow G$	None	
	7288	$\Lambda \Rightarrow G$	$M \Rightarrow V$	2868
C4	5901	$G \Rightarrow U$	$E \Rightarrow D$	2405
	6113	$A \Rightarrow U$	$H \Rightarrow L$	2476
C5	2890	$A \Rightarrow G$	$\kappa \Rightarrow \epsilon$	1402
C6	7209	A ⇒ G	None	

[0104]

In Table 4 and Fig. 11, "C1 to C6" represent replicon-replicating cell clones C1 to C6 having replicon RNA found to have mutations. "Nucleotide No." shows the corresponding nucleotide numbers within the nucleotide sequence of replicon RNA rSGREP-JFH1 (SEQ ID NO: 1). "Amino acid No." shows the corresponding amino acid numbers within the amino acid sequence encoded by the JFH-1 clone (SEQ ID NO: 4). The types of nucleotides and amino acids at mutation sites are described according to their general notations. As shown in Table 4, in clone C2, a nucleotide corresponding to nucleotide No. 4955 of SEQ ID NO: 1 on the replicon RNA mutated from C (cytosine) to U (uracil), which results in a mutation of an amino acid corresponding to amino acid No. 2090 of SEQ ID NO: 4 from A (alanine) to V (valine).

[0105]

Furthermore, mutation positions shown in Fig. 11 are shown with bar lines with the nucleotide numbers shown in Table 4. A thick bar line represents nonsynonymous substitution, and a thin bar line represents synonymous substitution.

[0106]

There were 2 clones having no nucleotide mutations at all that cause amino acid mutations. When Northern blot analysis was conducted for the 2 clones, it was shown that in these 2 clones, the quantity of replicon RNAs replicated was lower than those in the cell clones that had replicated replicon RNAs having a nucleotide mutation that causes an amino acid mutation. Hence, it was considered that the nucleotide mutation causing an amino acid mutation within the replicon RNA was an adaptive mutation for increasing the replication efficiency of the replicon RNA in Huh7 cells.

[0107]

[Example 6]

(J) Establishment of replicon-replicating cell clone using cells other than Huh7

cells

According to the method described in Example 1, rSGREP-JFH1 was transfected into some hepatic cancer cells other than Huh7 cells and non-liver-derived cells. The transfected cells were seeded into culture dishes and then cultured. Colony formation was observed and the number of colonies was counted. The cells used for transfection are as follows.

[0108]

- (1) HepG2 cells (representative hepatic cancer cells as well as Huh7 cells)
- (2) 1MY-N9 cells (established by Ito et al; fusion cells of HepG2 cells and human primary culture hepatic cells (Hepatology 2001, 34: 566-572))
- (3) HeLa cells (human cervical cancer-derived cells (Can Cer Res. 1952, 12: 264-265))
- (4) 293 cells (human fetal kidney-derived cells (Gen. Virol. 1977, 36: 59-72))

The results of transfection using HepG2 cells, IMY-N9 cells, HeLa cells or 293 cells, respectively, are shown in Fig. 12a to d. As shown in Fig. 12a to d, all HepG2 cells, IMY-N9 cells, HeLa cells, and 293 cells showed colony formation for rSGREP-JFH1-transfected cells.

[0110]

For the established cell clones, detection of the replicated replicon RNA, confirmation of the presence or the absence of the incorporation of the neomycin resistance gene into host genomic DNA, and confirmation of the expression of HCV protein were performed as described later, (L) and (M). The cell clones, for which the replication of the replicon in the cells had been confirmed, were regarded as replicon-replicating cell clones. Specifically, it was demonstrated that the use of rSGREP-JFH1 also enables the preparation of HCV replicon-replicating cells using hepatic cancer cells other than Huh7 cells and non-hepatic cells with which the production of HCV replicon-replicating cells had previously been unsuccessful (Blight et al., Science, (2000) 290; 1972-1974).

[0111]

(K) Detection of replicon RNA in replicon-replicating cells using cells other than Huh7 cells

Northern blot analysis was conducted according to a description of Molecular Cloning, A laboratory Manual, 2nd edition, J. Sambrook, E. F. Fritsch, T. Maniatis, Cold Spring Harbor Laboratory Press (1989). In accordance with the descriptions of the previous section (J), total RNA was extracted by the acidic phenol extraction method from each of the replicon-replicating cell clones that had been established by transfection of rSGREP-JFH1 into HepG2, IMY or HeLa cells respectively, and from pool clones of the replicon-replicating cells that had been established through transfection of rSGREP-JFH1 into 239 cells (prepared by collecting cell clones that had formed colonies from whole one dish and culturing them). Next, the total RNAs were analyzed by the Northern blot method using a pSGREP-JFH1-specific probe as a probe. As controls, total RNAs (lanes | and 17 in Fig. 13) extracted similarly from untransfected Huh7 cells and HepG2 cells, and RNA (lanes 2 and 3 in Fig. 13) prepared by adding 107 copies or 108 copies of the replicon RNA synthesized in vitro to total RNA extracted from Huh7 cells were used. As a result, RNA of approximately the same size of that of rSGREP-JFH1 was detected using a pSGREP-JFH1-specific probe (Fig. 13). Accordingly, it was confirmed that the replicon RNA derived from rSGREP-JFH1 that had been transfected at the beginning was replicated and proliferated within the cell clone. Furthermore, it was also revealed that the quantities of replicated replicon RNAs differed depending on cell type, and IMY cells were found to replicate the replicon RNA particularly efficiently. Moreover, it was revealed that the clones differed from each other in the quantity of the replicated replicon RNA.

[0112]

(L) Confirmation of the presence or the absence of the incorporation of the neomycin resistance gene into genomic DNA

For the thus established replicon RNA-replicating cell clone, PCR

amplification was performed using neomycin resistance gene-specific primers (sense primer, NEO-S3: 5'-AACAAGATGGATTGCACGCA-3' (SEQ ID NO: 29), antisense primer. NEO-R: 5'-CGTCAAGAAGGCGATAGAAG-3' (SEO ID NO: 30)) and the host cellular genomic DNA extracted from each of the cell clones as a template, in order to confirm that the resistance of each of the cell clones against G418 was not due to the incorporation of the neomycin resistance gene into the genome. The cell clones used herein were the cell clones Nos. 1, 5, 7, 8, 9, 10, 11, 12 and 13 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into HepG2 cells, and the cell clones Nos. 3, 4, 5, 6, 7, 8, 9, 10 and 11 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into IMY N9 cells. As a result, as shown in Fig. 14, in the nine examined cell clones obtained by introduction of rSGREP-JFHI into HepG2 cells, a positive clone showing the amplification of the neomycin resistance gene was not observed. In the 9 examined cell clones obtained by introduction of rSGREP-JFH1 into IMY N9 cells, a positive clone showing the amplification of the neomycin resistance gene was not observed.

[0113]

A similar examination was performed for cell clones obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into HeLa cells, and cell clones obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into 293 cells. Then, a positive clone showing the amplification of the neomycin resistance gene was not observed.

f01141

(M) Detection of HCV protein

Proteins were extracted from the established cell clones by a standard procedure, and then analyzed by SDS-PAGE and the Western blot method (Fig. 15). The cell clones examined in this case were the same as those used in the above section; the cell clones Nos. 1, 5, 7, 8, 9, 10, 11, 12 and 13 obtained by retransfection of rSGREP-JFH1-derived replicated replicate RNA into HepG2 cells,

and the cell clones Nos. 3, 4, 5, 6, 7, 8, 9, 10 and 11 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into IMY-N9 cells. Furthermore, according to a previous report (Lehmann et. al., Science, (1999)), the HCV RNA replicon-replicating cell clone prepared by introducing rSGREP-JFH1 into HuH7 was regarded as a positive control (Fig. 15, lane 4-1, C6). Moreover, a protein extracted from untransfected cells was used as a negative control (Fig. 15, lane N). Protein samples extracted from each cell clone were blotted onto PVDF membranes (Immobilon-P, Millipore), and then detection of NS3 protein encoded by the replicated replicon RNA was performed using anti-NS3-specific antibody (provided by Dr. Moradpour; Wolk B, et al, J. Virology, 2000, 74: 2293-2304). As shown in the upper section in Fig. 15, a protein of the same size as that of the positive control was detected in the cell clones Nos. 1, 5, 7, 8, 9, 10, 11, 12 and 13 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA, and in the cell clones Nos. 3, 4, 5, 6, 7, 8, 9, 10 and 11 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into IMY N9 cells.

Moreover, by the use of the serum of a hepatitis C patient as an antibody, the confirmation of the expression of NS5a protein from the replicon RNA was performed for each cell clone that had been confirmed above to express NS3 protein. In this experiment, examination was performed in a manner similar to that in the case of the expression of NS3 protein, but using an antibody instead of the serum of the patient. As a result, as shown in the lower section in Fig. 15, a protein of the same size as that of the positive control was detected in the cell clones Nos. 1, 5, 7, 8, 9, 10, 11, 12 and 13 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA, and the cell clones Nos. 3, 4, 5, 6, 7, 8, 9, 10 and 11 obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into IMY N9 cells.

[0116]

When similar examination was performed for the cell clones obtained by

retransfection of rSGREP-JFH1-derived replicated replicon RNA into HeLa cells, and the cell clones obtained by retransfection of rSGREP-JFH1-derived replicated replicon RNA into 293 cells, the expression of NS3 and that of NS5a proteins could be confirmed.

[0117]

As described above, it was confirmed that the replicon RNA was replicated in the cell clones that had been established through transfection of the replicon RNA.

[0118]

[Example 7]

(N) Analysis of adaptive mutation

According to the descriptions of Example 3, total RNAs obtained from the replicon-replicating cell clones established through the transfection of rSGREP-JFH1 into HepG2 and HeLa cells were re-transfected into another cells of the each cell line, respectively, so that 14 cell clones were established for HepG2 cells and 8 cell clones were established for HeLa cells. From each of these cell clones, total RNA was extracted by a standard procedure. cDNA corresponding to the replicon RNA was synthesized using the total RNA as a template, reverse transcriptase Superscript II (Invitrogen) and a primer 9641R-IH (5'-GCACTCTCTGCAGTCATGCGGCTCACGGAC-3' (SEQ ID NO: 31)). composition of a reaction solution for the synthesis of cDNA by reverse transcription reaction is as shown below.

[0119]

Fluid Volume (µl) Composition of Reaction Solution

Composition of Reaction Solution	Fluid Volume
5x 1st strand Buffer	4
	5
2 mM dNTP	1
0.1 M DTT	
9651R-IH primer (100 μM)	

DW (distilled water)	6.5
Sample RNA (2 mg/mL)	1
RNAsin (Promega)(40 U/μL)	0.5
Superscript II RT (Invitrogen)	
Total	20 μl

[0120]

In cDNA synthesis reaction, the above reagents other than RNAsin and Superscript II were mixed to prepare a first reaction solution. The first reaction solution was heated at 90°C for 3 minutes, and then cooled on ice. Subsequently, RNAsin and Superscript II were added to the reaction solution, and then the solution was allowed to react at 42°C for 1 hour, followed by another reaction at 70°C for 15 minutes.

[0121]

Furthermore, PCR amplification was performed using the thus obtained cDNA together with five primer sets by the following procedures, so that DNA amplification fragments covering almost all the regions of replicon RNA were obtained. The primer sets used and regions amplified by each primer set are shown in Table 5 and Table 6 below.

[0122] [Table 5]

Designation of amplified fragment		Primer set	Amplified region
ampirited Haginer	Primer 1	Primer 2	
A	42S-IH	433R-neo	41-470
3	C/S17ssp	4680R-IH	28-3026
	4534S-IH	7279R-IH	2280-5625
2	7198S-IH	9367R-IH	5544-7713
D E	9247S-NF	9576R-NF	7597-7966

In this table, an amplified region is represented by nucleotide numbers in

rSGREP-JFH1 (SEQ ID NO: 1) that the region corresponds to.
[0123]
[Table 6]

[rable of		
Primer	Nucleotide Sequence (5' to 3')	SEQ ID NO:
Designation		ano ID NO. 14
43S-IH	CCCCTGTGAGGAACTACTGTCTTCACGC	SEQ ID NO: 14
C/S17ssp	CCGGGAGAGCCATAGTGGTCTGCG	SEQ ID NO: 15
/ 1	CCACTCAAAGAAAAAGTGTGACGAGCTCGC	SEQ ID NO: 16
4534S-IH	GGCTTGGGCACGGCCTGA	SEQ ID NO: 17
7198S-IH		SEO ID NO: 18
9247S-NF	GCGGTGAAGACCAAGCTCAAACTCACTCCA	
433R-neo	AGAACCTGCGTGCAATCCATC	SEQ ID NO: 19
4680R-IH	CCCGTCATGAGGGCGTCGGTGGC	SEQ ID NO: 20
	ACCAGCAACGGTGGGCGGTTGGTAATC	SEQ ID NO: 21
7279R-IH	***	SEQ ID NO: 22
9367R-IH	GGAACGCGACACGCTGTG	-
9576R-NF	AGCTAGCCGTGACTAGGGCTAAGATGGAGC	SEO ID NO: 23
9576R-NF	AGCTAGCCGTGACTAGGGCTAAGACCGG	

[0124]

The composition of a reaction solution in this PCR reaction is as follows.

The composition of a reaction solution			
Composition of Reaction Solution	Fluid Volume (µl)		
Primer 1 (10 µM)	1.0		
Primer 2 (10 µM)	1.0		
2.5 mM dNTPs	5.0		
10x LA Buffer	5.0		
MgCl ₂ (25 mM)	5.0		
LA Taq (TAKARA) (5 U/μl)	0.3		
DW (distilled water)	30.7		
Template cDNA	2.0		
Total	50 μl		

[0125]

[0126]

In addition, PCR reaction conditions are as follows: 95°C for 2 minutes; 35 cycles of 98°C for 10 seconds and then 68°C for 8 minutes; followed by 72°C for 7 minutes, after which the temperature is kept at 4°C.

The nucleotide sequence of each PCR product obtained as described above was determined, and then the RNA sequence corresponding to the DNA sequence was compared with the sequence of rSGREP-JFH1. The results are shown in Table 7 and Table 8.

[Table 7]

Analysis of adaptive mutation of JFH-1 replicon in HepG2 cells

Clone	Mutat	on site	Mutation			
	Nucleotide No.	Amino acid No.	Nucleotide	Amino acid		
HepIH1	6826	2714	C⇒A	Q⇒K		
HepIH3	6887	2734	C⇔A	T⇒N		
HepIH5	6887		U⇒A	None		
HepIH8	6580	2632	U⇒A	$S \Rightarrow T$		
	7159	2825	U⇒C	Y⇒H		
Hep1H9	3342		A⇒G	None		
	3594		C⇒A	None		
	7230	2848	U⇒A	N⇒K		
HepIH10	5052		U⇒C	None		
	6943	2753	C⇒A	$P \Rightarrow T$		
HepIH12	None					
HepIH13	4302		C⇒U	None		
	5687	2334	G⇒A	G⇒D		
	6110	2475	A⇒G	Y⇒C		

[0127]

As shown in Table 7, in the case of HepG2 cells, a total of 13 nucleotide mutations were observed in 8 cell clones, and 8 of these mutations were nonsynonymous substitutions that cause amino acid mutations. Types of these

mutations are shown in Table 8. On the other hand, in the case of HeLa cells, a total of 7 nucleotide mutations were observed in 3 cell clones, and 5 of these mutations were nonsynonymous substitutions that cause amino acid mutations. Types of these mutations are shown in Table 8.

[0128]

[Table 8]

Analysis of adaptive mutation of JFH-1 replicon in HeLa cells

Clone	Mutat	ion site	Mutation			
	Nucleotide No.	Amino acid No.	Nucleotide	Amino acid		
HeLaH1	None					
HeLaIH2	5550	2272	U⇒C	S⇒P		
	6252		A⇒G	None		
	7182		U⇒C	None		
	7217	2844	A⇒G	H⇒R		
HeLaIH5	3643	1653	A⇒G	M⇒V		
	5851	2389	G⇒A	A⇒T		
	5914	2410	G⇒A	$E \Rightarrow K$		

[0129]

In Tables 7 and 8, "HepIH No." represents clone numbers of repliconreplicating cell clones that have replicon RNA and have been cloned using HepG2
cells. "Nucleotide No." shows the corresponding nucleotide number in the
nucleotide sequence (SEQ ID NO: 1) of replicon RNA rSGREP-JFH1. "Amino
acid No." shows the corresponding amino acid number in the amino acid sequence
(SEQ ID NO: 4) encoded by the JFH-1 clone. The types of nucleotides and
amino acids at mutation sites are described according to their general notations.
As shown in Table 7, for example, in clone HepIH1, a nucleotide corresponding to
nucleotide No. 6826 of SEQ ID NO: on the replicon RNA mutated from C to A, so
that an amino acid corresponding to amino acid No. 2714 of SEQ ID NO: mutated
from Q to E. Similarly, in Table 8, "HeLaIH No." represents numbers of
replicon-replicating cell clones that have replicon RNA and have been cloned

using HeLa cells.

[0130]

In addition, when Northern blot analysis was conducted for clones having no nucleotide mutations at all that cause amino acid mutations, it was shown that the quantity of replicon RNA replicated by the clones was lower than that of a cell clone replicating replicon RNA having a nucleotide mutation that causes an amino acid mutation. Hence, it was concluded that the nucleotide mutation in replicon RNA inducing an amino acid mutation was an adaptive mutation for increasing the replication efficiency of replicon RNA in cells.

[Industrial Applicability]

[0131]

The replicon-replicating cells according to the present invention can be utilized as a culture system for the continuous production of HCV genotype 2aderived RNA and HCV protein. Moreover, the replicon-replicating cells according to the present invention are useful as a test system for screening for various substances affecting the replication of HCV and/or the translation into HCV protein.

[Brief Description of Drawings]

[0132]

[Fig. 1] Fig. 1 is a schematic view showing procedures for constructing a template DNA for preparing the HCV-RNA replicon according to the present invention. The upper section of Fig. 1 shows the structure of the region within pJFH1 or pJCH1, with the viral genome inserted into it. The lower section of Fig. 1 shows the structure of the region within plasmid DNA pSGREP-JFH1 or pSGREP-JCH1, with the viral genome inserted into it, that had been constructed by substituting a part of viral genome-inserted region of pJFH1 or pJCH1 with a DNA fragment containing a neomycin resistance gene and EMCV IRES. Symbols in Fig. 1 are as described below. T7, T7 RNA promoter; G, dGTP that was inserted upstream of the 5' end of the inserted DNA derived from JFH-1 or JCH-1

and downstream of the 3' end of T7 RNA promoter sequence; 5' NTR, 5' untranslated region; Core, core protein; and 3' NTR, 3' untranslated region. E1 and E2 represent envelope proteins. NS2, NS3, NS4A, NS4B, NS5A and NS5B represent non-structural proteins. Age I, Cla I and Xba I represent cleavage sites of restriction enzymes Age I, Cla I and Xba I, respectively. GDD, the position of amino acid motif GDD corresponding to the active center of NS5B protein; neo, neomycin resistance gene; and EMCV IRES, internal ribosome entry site of encephalomyocarditis virus (EMCV IRES).

[Fig. 2A] Fig. 2A shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2B] Fig. 2B shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2C] Fig. 2C shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2D] Fig. 2D shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2E] Fig. 2E shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 2F] Fig. 2F shows the nucleotide sequence of rSGREP-JFH1.

[Fig. 3A] Fig. 3A shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3B] Fig. 3B shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3C] Fig. 3C shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3B] Fig. 3B shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 3E] Fig. 3F shows the nucleotide sequence of rSGREP-JCH1.

[Fig. 4] Fig. 4 shows photographs showing the colony formation of Huh7 cells to which rSGREP-JFHI, rSGREP-JFHI/GND and rSGREP-JFHI/dGDD was transfected, respectively. The amount of each of three transfected RNAs in the upper section was 100 ng and that of three transfected RNAs in the lower section was 300 ng.

[Fig. 5] Fig. 5 shows photographs showing colony formation of Huh7 cells to which rSGREP-JFH1 and rSGREP-JCH1 respectively had been transfected when the concentration of G418 was 0.5 mg/ml of the medium. The amount of each of these RNAs transfected was 100 ng.

[Fig. 6] Fig. 6 shows photographs showing the effect of Mung Bean Nuclease treatment conducted on the colony-forming ability of the transfected cells. The amount of rSGREP-JFH1 RNA transfected was 100 ng for both cases. The concentration of G418 was 1.0 mg/ml in both media.

[Fig. 7] Fig. 7 shows photographs showing colony formation when total cellular RNA derived from the replicon-replicating cell clone, which had been established by transfection of rSGREP-JFHI, was retransfected to another Huh7 cells. The photograph on the left shows that the formation of 96 colonies was observed as a result, when using the total cellular RNA derived from the replicon-replicating cell clone No. 6. The photograph on the right shows that the formation of 77 colonies was observed as a result, when using the total cellular RNA derived from the pool clones. In both cases, RNA was retransfected in an amount containing 1×10^7 copies of the replicon RNA.

[Fig. 8] Fig. 8 shows photographs showing the results of detecting by the Northern blot method using an rSGREP-JFH1-specific probe for the total RNA derived from a cell clone that had been obtained by retransfecting the total cellular RNA (derived from the replicon-replicating cell clone established by transfection of rSGREP-JFH1) into another Huh7 cells. Explanation of the lanes is as follows. 108 represents sample prepared by adding 108 copies of the replicon RNA synthesized in vitro to total RNA extracted from Huh7 cells. 107 represents sample prepared by adding 107 copies of the replicon RNA synthesized in vitro to total RNA extracted from Huh7 cells. Huh7, total RNA extracted from untransfected Huh7 cells; pool clone, total RNA extracted from the pool clones; and 1-11, total RNA extracted from each of cell clones Nos. 1 to 11. "Replicon RNA" represents the electrophoresed position of a molecular weight marker indicating the size of rSGREP-JFH1, "28S" represents the same of a ribosomal RNA marker indicating the size of molecular weight of 4.5 kb, and "18S" represents the same of a ribosomal RNA marker indicating the size of molecular weight of 1.9 kb.

[Fig. 9] Fig. 9 shows photographs showing the presence or the absence of the incorporation of a neomycin resistance gene into the genomic DNA of a host cell in the cell clone to which rSGREP-JFH1- or rSGREP-JCH1-derived replicated replicon RNA was retransfected. Explanation of the lanes in the photograph on the left is as follows. M, DNA molecular weight marker; 1-8, rSGREP-JFH1-derived cell clones Nos. 1 to 8; N, untransfected Huh7 cells; and P, positive control (PCR amplification product of the neomycin resistance gene). Furthermore, explanation of the lanes in the photograph on the right is as follows. M, DNA molecular weight marker; and 1-6, rSGREP-JCH1-derived cell clones Nos. 1 to 6.

[Fig. 10] Fig. 10 shows photographs showing the results of detecting NS3 protein expressed in the cell clone that was retransfected with rSGREP-JFH1- or rSGREP-JCH1-derived replicated replicon RNA. Lanes 1 to 8 of the photograph on the left represent rSGREP-JFH1-derived cell clones Nos. 1 to 8. Lanes 1-6 of the photograph on the right represent rSGREP-JCH1-derived cell clones Nos. 1 to 6. Lane P of the photograph on the right represents NS3 protein (positive control) and N represents protein extracted from untransfected Huh7 cells (negative control).

[Fig. 11] Fig. 11 shows the positions of nucleotide mutations in replicon RNAs obtained from 21 cell clones that were established through the retransfection of rSGREP-JFH1-derived replicated replicon RNA into Huh7 cells. Mutation positions are indicated using bar lines shown with nucleotide numbers listed in Table 4. A thick bar line denotes nonsynonymous substitution and a thin bar line denotes synonymous substitution.

[Fig. 12] Fig. 12 shows photographs showing the results of transfection with rSGREP-JFH1 using 1) HepG2 cells; 2) IMY-N9 cells; 3) 293 cells; or 4) HeLa cells

[Fig. 13] Fig. 13 shows photographs showing the results of performing Northern blotting for replicon-replicating cell clones. [Fig. 14] Fig. 14 shows photographs showing the results of electrophoresis performed for confirming the incorporation of the neomycin resistance gene into genomic DNA.

[Fig. 15] Fig. 15 shows photographs showing the results of analyzing by the Western blot method proteins derived from the replicon-replicating cell clones. [Sequence Listing Free Text]

[0133]

SEQ ID NO: 1. Explanation of artificial sequence: replicon

SEO ID NO: 2. Explanation of artificial sequence: replicon

SEO ID NO: 7. Explanation of artificial sequence: replicon

SEQ ID NOS: 8 to 12. Explanation of artificial sequences: synthetic RNA

SEQ ID NOS: 13 to 41. Explanation of artificial sequences: synthetic DNA

[Sequence Listing]

SEQUENCE LISTING

<110> Toray Industries Inc.

Tokyo Metropolitan Organization for Medical Research Johannes Gutenberg-Universitaet Mainz

<120> Establishment of the genotype 2a Hepatitis C virus subgenomic replicon

<130> P03-0852

<140>

<141>

<150> JP 2003-148242 <151> 2003-05-26

<160> 41

```
<170> PatentIn Ver. 2.1
<210> 1
<211> 8024
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: replicon
<220>
<223> Inventor: Wakita, Takaji
     Inventor: Kato, Takanobu
     Inventor: Date, Tomoko
<400> 1
accuqccccu aauaqqqqcq acacuccqcc auqaaucacu ccccuquqaq qaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
ccccuccc qqaqaccau aququcuqc qqaaccqquq aquacaccqq aauuqccqqq 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
egeuugegag ugeeeeggga ggueueguag acegugeace augageacaa auccuaaace 360
ucaaagasaa accaaaagaa acaccaaccg ucqcccaaug auugaacaag auggauugca 420
cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480
aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggcgcc cgguucuuuu 540
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600
guggouggoo acgacgggcg uuccuugcgc agcugugouc gacguuguca cugaagcggg 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccugcogag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780
ggcuaccuge ecauucgace accaagegaa acaucgeauc gagegageac guacucggau 840
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900
```

cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggegaugee ugeuugeega auaucauggu ggaaaaugge egeuuuueug gauucauega 1020 cuguggcogg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggeggeg aaugggeuga cegeuuceue gugeuuuaeg guauegeege 1140 uccegauucg cagegeaucg ceuucuaucg ceuucuugae gaguucuucu gaguuuaaae 1200 ccucucccuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260 cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320 aaccuggeee ugucuucuug acgageauuc cuaggggucu uuccecucuc gecaaaggaa 1380 ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440 caacgucugu agcgacccuu ugcaggcagc ggaacccccc accuggcgac aggugccucu 1500 geggecaaaa gecaegugua uaagauacae cugcaaagge ggeacaaece cagugecaeg 1560 uugugaguug gauaguugug gaaagaguca aauggeucuc cucaagegua uucaacaagg 1620 ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680 caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggcccccg aaccacgggg 1740 acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800 caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860 caggoogggg aaguccaaau couguccaca gucucucagu couucoucgg aacaaccauc 1920 ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980 gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gcccagcccc 2040 ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100 cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160 ccgagaccca uuucgaccuu gaaggggucc ucgggggggc cggugcucug cccuaggggc 2220 cacquequug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280 uucauccccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340 acgccaccgg cugugeccca gaccuaucag gucggguacu ugcaugcucc aacuggcagu 2400 ggaaagagca ccaagguccc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460 aaccccucgg uagcugccac ccugggguuu ggggcguacc uauccaaggc acauggcauc 2520 aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580 acauauggca aauuucucgc cgaugggggc ugogcuagcg gogccuauga caucaucaua 2640 ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700 caagcagaga cagccggggu cagacuaacu gugcuggcua cggccacacc ccccggguca 2760 gugacaaccc cccaucccga uauagaagag guaggccucg ggcgggaggg ugagaucccc 2820

uucuauggga gggcgauucc ccuauccugc aucaagggag ggagacaccu gauuuucugc 2880 cacucaaaga aaaaguguga cgagcucgeg geggeeeuuc ggggcauggg cuugaaugec 2940 guggcauacu auagaggguu ggacgucucc auaauaccag cucagggaga uguggugguc 3000 guogecaceg aegoccucau gaegggguae acuggagacu uugacucegu gauegacuge 3060 aauguagegg ucacccaage uguegacuuc agccuggace ccaccuucae uauaaccaca 3120 caqacuquee cacaagaege ugucucaege agucagegee gegggegeae agguagagga 3180 agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240 guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300 accaccquea ggcuuagage guauuucaac acgcccggcc uacccgugug ucaagaccau 3360 cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucucc 3420 caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480 gecagageca aggececuce ceequecugg gacgecaugu ggaagugecu ggccegacue 3540 aagccuacgc uugcgggccc cacaccucuc cuguaccguu ugggcccuau uaccaaugag 3600 gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660 gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720 cuggegacug gaugeguuuc caucaucgge egeuugeacg ucaaccageg aguegueguu 3780 gcgccggaua aggagguccu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840 geggeueuca uegaagaggg geageggaua geegagaugu ugaagueeaa gauceaagge 3900 uugcugcage aggeeucuaa geaggeeeag gacauacaae eegeuaugea ggeuucaugg 3960 cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020 cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080 geogeecuca ccagueeguu guegaccagu accaccauce uucucaacau caugggagge 4140 ugguuagegu eccagauege accaeeegeg ggggeeaeeg geuuuguegu caguggeeug 4200 gugggggcug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260 uauggugegg geauuueggg ggeceuegue geauucaaga ucaugueugg egagaageee 4320 ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380 ggggucaucu gegeggeeau ucugegeege cacgugggae egggggaggg egegguccaa 4440 uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500 gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560 agccuacuca gaagacucca caauuggaua acugaggacu gccccauccc augcuccgga 4620 uccuggeucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680 cugaccucua aauuguuccc caagcugccc ggccuccccu ucaucucuug ucaaaagggg 4740

uacaagggug ugugggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800 aucucuggca auguccgccu gggcucuaug aggaucacag ggccuanaac cugcaugaac 4860 accuggeagg ggaccuuucc uaucaauugc uacacggagg gccagugege geegaaaccc 4920 cccacgaacu acaagacegc caucuggagg guggeggccu cggaguacgc ggaggugacg 4980 cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aauuccuugc 5040 caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100 cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggeu uaauuecuau 5160 gcuguegggu eccagcuuce cugugaaccu gagcccgacg cagacguauu gagguecaug 5220 cuaacagauc egececacau caeggeggag acugeggege ggegeuugge aeggggauca 5280 ccuccaucug aggcgagcuc cucagugagc cagcuaucag caccgucgcu gcgggccacc 5340 ugcaccacce acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400 ggeggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460 auggoogagg aagagagoga couugageee ucaauaccau oggagugcau goucccoagg 5520 agoggguuuc caogggccuu acoggcuugg gcaeggeeug acuacaacce gccgcucgug 5580 gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucccccc 5640 cccaaqaaqq ccccqacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700 accauaucaq aagccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760 ggugaugcag gcucquccac gggggegggc gcegeegaau ceggegguec gacquecccu 5820 ggugageegg eccecucaga gacagguuce gccuccucua ugccccccu egagggggag 5880 ccuggagauc eggaccugga gucugaucag guagagcuuc aaccucccc ccaggggggg 5940 gggguagcuc cegguueggg eueggggueu uggueuacuu geueegagga ggaegauace 6000 accgugugeu geuccaugue auacuccugg accggggeue uaauaacuce cuguagecee 6060 gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120 guguacugua caacaucaaa gagegecuca cagagggcua aaaagguaac uuuugacagg 6180 acgcaagugc ucgacgccca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240 aaggucagog caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 geuegeeuca ueguunaece ugaceuegge gueegggueu gegagaaaau ggeeeucuau 6540 gacauuacac aaaagcuucc ucaggeggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660

uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugccc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugec gegecagegg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau gcugguaugc 6960 ggcgaugace uaguagucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020 agagecunea eggaggeeau gaceagguae ucugeeccue cuggugauce ecceagaceg 7080 qaauauqacc ugqagcuaau aacauccugu uccucaaaug ugucuguggc guugggcccg 7140 oggggeegee geagauacua ceugaceaga gacecaacea cuccacuege cogggeugee 7200 uqqqaaacaq uuaqacacuc cccuaucaau ucauqqcugg gaaacaucau ccaguaugcu 7260 ccaaccauau ggguucgcau gguccuaaug acacacuucu ucuccauucu caugguccaa 7320 gacacccugg accagaaccu caacuuugag auguauggau caguauacuc cgugaauccu 7380 uuggaccuuc cagccauaau ugagagguua cacgggcuug acgccuuuuc uaugcacaca 7440 uacucucacc acgaacugac geggguggcu ucageccuca gaaaacuugg ggegccaccc 7500 cucagggugu ggaagagucg ggcucgcgca gucagggcgu cccucaucuc ccguggaggg 7560 aaagcggccg uuugcggccg auaucucuuc aauugggcgg ugaagaccaa gcucaaacuc 7620 acuccauuqc cgqaggcgcg ccuacuggac uuauccaguu gguucaccgu cggcgccggc 7680 qqqqqqaca uuuuucacag cgugucgcgc gcccgacccc gcucauuacu cuucggccua 7740 cuccuacuuu ucguaggggu aggccucuuc cuacuccccg cucgguagag cggcacacac 7800 иаддиасаси ссанадсная сидинссини инпининии инпининии инпининии 7860 uuuuuuuuu cuuuuuuuu uuuucccuc uuucuucccu ucucaucuua uucuacuuuc 7920 uuucuuggug gcuccaucuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980 augacugcag agagugccgu aacuggucuc ucugcagauc augu 8024

```
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: replicon
```

<210> 2 <211> 8024

<400> 2

acceqcecu aauaggggeg acacucegee augaaucacu ceccugugag gaacuacugu 60 cuucacgcag aaagcgucua gccauggegu uaguaugagu gucguacagc cuccaggeec 120 cccccucccg ggagagccau aguggucuge ggaaccggug aguacaccgg aauugccggg 180 aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 ugcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa aucccaaacc 360 ucaaagaaa accaaaagaa acacuaaccg ucgcccaaug auugaacaag auggauugca 420 cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480 aaucggcugc ucugaugccg ccguguuccg gcugucagcg cagggggggcc cgguucuuuu 540 ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcgggcuauc 600 guggouggce acgaeggcg uuccuugcge ageugugeue gaeguuguea eugaageggg 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugeegag aaaguaucca ucauggeuga ugcaaugegg eggeugeaua egeuugauce 780 ggouaccugo ccauucgaco accaagogaa acaucgcauc gagogagcac guacucggau 840 ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900 cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020 cuquqqcqq cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140 uccogauucg cagogcaucg couucuaucg couucuugac gaguucuucu gaguuuaaac 1200 ccucucccue ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260 cquuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320 aaccuggeee ugucuucuug acgagcauuc cuaggggucu uuccccucuc gocaaaggaa 1380 ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440 caacgucugu agegacccuu ugcaggcagc ggaacccccc accuggcgac aggugccucu 1500 gcggccaaaa gccacgugua uaagauacac cugcaaagge ggcacaacce cagugccacg 1560 uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagegua uucaacaagg 1620 ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680 caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggccccccg aaccacgggg 1740 acqugguuuu ccuuugaaaa acacgauaau accauggccc ccaucaccgc uuacgcccag 1800

cagacacgag gucucuuggg cucuauagug gugagcauga cggggcguga caagacagaa 1860 caggeegggg aggueeaagu eeugueeaca gucacucagu eeuuceuegg aacauceauu 1920 ucgggggucu uauggacugu uuaccacgga gcuggcaaca agacacuagc cggcucgcgg 1980 ggcccgguca cgcagaugua cucgagcgcc gagggggacu uggucgggug gcccagcccu 2040 ccugggacca aaucuuugga gccguguacg uguggagcgg ucgaccugua uuuggucacg 2100 cggaacgcug augucauccc ggcucgaaga cgcggggaca agcggggagc gcugcucucc 2160 ccgagacccc uuucgaccuu gaaggggucc ucggggggac cugugcuuug cccuaggggc 2220 cacgougueg gaaucuuceg ggeageugug ugcucucggg guguggeuaa guccauagau 2280 uucauccccg uugagacgcu cgacaucguc acgcggucuc ccaccuuuag ugacaacagc 2340 acaccaccag cugugcccca gaccuaucag gugggguacu ugcacgcccc cacuggcagu 2400 ggaaaaagca ccaagguccc cgucgcguac gccgcccagg gguauaaagu gcuggugcuc 2460 aaucccucgg uggcugccac ccugggauuu ggggcguacu uguccaaggc acauggcauc 2520 aaccccaaca uuaggacugg agucagaacu gugacgaceg gggagcccau uacauacucc 2580 acquauggua aauuccucgc cgaugggggc ugcgcaggeg gegccuauga caucaucaua 2640 ugcgaugaau gccacucugu ggaugcuacc acuauucucg gcaucgggac aguccuugac 2700 caagcagaga cagccggggu caggcuaacu guacuggcca cggccacgcc ccccgggucg 2760 gugacaaccc cccaucccaa uauagaggag guagcccucg gacaggaggg ugagaucccc 2820 uucuauggga gggcguuucc ccugucuuac aucaagggag ggaggcacuu gauuuucugc 2880 cacucaaaga aaaaguguga cgagcucgca acggcccuuc ggggcauggg cuugaacgcu 2940 guggeauauu acagaggguu ggacgucucc auaauaccaa cucaaggaga uguggugguc 3000 guugecaceg aegeceucau gaegggguau aeuggagaeu uugaeueegu gauegaeuge 3060 aacguagegg ucacccaggc cguagacuuc agccuggacc ccaccuucac uauaaccaca 3120 cagacuguec egeaagaege ugucucaegu agucagegee gagggegeae ggguagagga 3180 agacugggca uuuauaggua uguuuccacu ggugagcgag ccucaggaau guuugacagu 3240 guaguacucu gugagugcua cgacgcagga gcugcuuggu augagcucuc accaguggag 3300 acgaceguca ggeucaggge guauuucaac acgccuggcu ugccugugug ccaggaccac 3360 cagacaaagc agucggggga aaauuucgca uacuuaguag ccuaucaggc cacagugugc 3480 aagcccacgc uugugggccc uacaccucuc cuguaccguu ugggcucugu uaccaacgag 3600 gucacccuua cacaccccgu gacaaaauac aucgccacau gcaugcaagc ugaccucgag 3660 gucaugacca gcacgugggu ccuggcuggg ggagucuuag cagccgucgc cgcguauugc 3720

uuagegaeeg gguguguuue caucauugge eguuuacaca ucaaccageg ageuguegue 3780 gcuccggaca aggagguccu cuaugaggcu uuugaugaga uggaggaaug ugccuccaga 3840 geggeueuce uugaagaggg geageggaua geegagauge ugaagueeaa gaueeaagge 3900 uuauugcago aagocucuaa acaggoocag gacauacaac cogcugugca agouucgugg 3960 cccaagaugg agcaauucug ggccaaacau auguggaacu ucauaagcgg cauucaguac 4020 cucgcaggac ugucaacacu gccagggaac ccugcugugg cuuccaugau ggcauucagc 4080 gccgcccuca ccaguccguu gucaacuagc accaccaucc uucuuaacau ucuggggggc 4140 uggeuggegu cccaaauugc gccaccegeg ggggccacug gcuuuguugu caguggecug 4200 gugggagcug cuguuggcag cauaggcuug gguaaagugc ugguggacau ccuggcaggg 4260 uauggugegg gcauuucggg ggcccucguc gcguuuaaga ucaugucugg cgagaagccc 4320 uccauggagg augucaucaa cuugcugccu gggauucugu cuccaggugc ucugguggug 4380 ggagucaucu gogoggccau ucugogccgc caugugggac cgggggaagg cgcgguccaa 4440 uggaugaaca ggcuuaucgc cuucgcuucc agaggaaacc acgucgcccc uacucacuac 4500 gugacggagu cggaugcguc gcagcguguc acccaacugc uuggcucucu cacuauaacu 4560 agucuacuca ggagacuuca caacuggauc acugaggauu gccccauccc augcgccggc 4620 ucguggcucc gcgaugugug ggacuggguc uguaccaucc uaacagacuu uaagaacugg 4680 cugaccucca agouguuccc aaagaugccu ggccuccccu uuaucucuug ccaaaagggg 4740 uacaagggcg ugugggccgg cacuggcauc augaccacac gaugccccug cggcgccaac 4800 aucucuggca acguccgcuu gggcucuaug agaaucacag gacccaaaac cugcaugaac 4860 accuggeagg ggaccuuucc uaucaauugu uauacagaag gccagugcuu gccgaaaccc 4920 geguuaaacu ucaagacege caucuggaga guggeggeeu cagaguaege ggaagugaeg 4980 cagcacggau cauaugccua uauaacaggg cugaccacug acaacuuaaa agucccuugc 5040 caacuccccu cuccagaguu uuucucuugg guggacggag uacaaaucca uagguccgcc 5100 cccacaccaa agccguuuuu ccgggaugag gucucguuca gcguugggcu caauucauuu 5160 gueguegggu cucageuuce cugugaeccu gagecegaca cugagguagu gauguecaug 5220 cuaacagacc caucccauau cacggcggag gcugcagcgc ggcguuuagc gcggggguca 5280 ccccaucug aggcaagcuc cucagegage cageuguegg egecauegeu gegagecaee 5340 ugcaccaccc acgguaggac cuaugaugu gacauggugg augccaaccu guucaugggg 5400 ggcggcguga uucggauaga gucugagucc aaaguggucg uucuggacuc ccucgacuca 5460 augaccgagg aagagggcga ccuugagccu ucaguaccau cggaguauau gcuccccagg 5520 aagagguuco caceggeeuu aceggeuugg gegeggeeug auuacaacee acegeuugug 5580 gaaucgugga agaggccaga uuaccaacca cccacuguug cgggcugugc ucucccccc 5640 cccaaaaaga ccccgacgcc uccuccaagg agacgccgga cagugggucu gagcgagagc 5700 accauaggag augocoucca acagouggee aucaagueeu uuggocagoo coccocaago 5760 ggegauucag gccuuuccac gggggeggac gccgccgacu ccggcgaucg gacacccccu 5820 gacgaguugg cucuuucgga gacagguucu accuccucca ugcccccccu cgagggggag 5880 ccuggggacc cagaccugga gccugagcag guagagcuuc aaccuccucc ccaggggggg 5940 gaggeageue eeggeuegga euegggguee uggueuaeuu geueegagga ggaugaeuee 6000 guegugugeu geuccaugue auauuccugg accggggeuc uaauaacucc uuguagcccc 6060 gaagaggaaa aguugccaau uaacuccuug agcaacucgc uguugcgaua ccauaacaag 6120 guauacugua cuacaucaaa gagugccuca cuaagggcua aaaagguaac uuuugauagg 6180 augcaagugc ucgacgccua uuaugauuca gucuuaaagg acaucaagcu agcggccucc 6240 aaggucageg caaggcuccu caccuuagag gaggcgugcc aauugacccc accccacucu 6300 gcaagaucca aguauggguu uggggcuaag gagguccgca gcuuguccgg gagggccguc 6360 aaccacauca aguccgugug gaaggaccuc uuggaagacu cacaaacacc aauuccuaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaaaaaacca 6480 gcucgccuua ucguuuaccc ugaccucggc gucagggucu gcgagaagau ggcccuuuau 6540 gaugucacac aaaagcuucc ucaggcggug augggggcuu cuuauggcuu ccaguacucc 6600 cccgcucagc ggguggaguu ucucuugaag gcaugggcgg aaaagagaga cccuaugggu 6660 uuuucguaug auaccogaug cuuugacuca accgucacug agagagacau caggacugag 6720 gaguccauau accaggocug cuccuuacco gaggaggoco gaacugocau acacuogcug 6780 acugagagac ucuauguggg agggcccaug uucaacagca agggccaguc cugcggguac 6840 aggeguugec gegeeagegg ggugeuuace acuaguaugg ggaacaccau cacaugcuau 6900 guasaagccc uagcggcuug caaggcugcg gggauaauug cgcccacgau gcugguaugc 6960 ggcgacgacu uggucgucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020 agagecunca eggaggenau gaccagguau ueugeceeue euggugaeee ecceagaeeg 7080 gaauaugacc uggagcuaau aacaucuugu uccucaaacg ugucuguggc acuuggccca 7140 cagggeegee geagauacua ceugaceaga gaeeceacea cuucaauuge cegggeugee 7200 uqqqaaacaq uuaqacacuc cccugucaau ucauqqcuqq gaaacaucau ccaguacgcu 7260 ccaaccauau ggguucgcau gguccugaug acacacuucu ucuccauucu cauggcccag 7320 gacacccuag accagaaccu uaacuuugaa auguacggau cgguguacuc cgugaguccu 7380 cuggaccucc cagecauaau ugaaagguua cacgggcuug acgccuucuc ucugcacaca 7440 uacacucece acgaacugae geggguggeu ucageecuca gaaaacuugg ggegceacce 7500 cucagagegu ggaagagueg ggegegugea guuagggegu cecucaucuc cegugggggg 7560
 aggggggggg
 uuugcgguog
 guaccucuuc
 acuuggggg
 ugauagacaa
 guacaacaa
 7620

 acuccuuug
 cggaggcaca
 cuccuuggau
 uuguccaguu
 gguuuaccag
 cggcgcggg
 7860

 gggggggaca
 uuuaucaaag
 cgugucgcgu
 cccaacacc
 gcuauuaccag
 ccuuauacuu
 agccucuu
 cucgaacaaag
 gggcacaca
 7800

 uuuuuuuuuu
 cuuuuuuuuu
 cuuuuuuuuu
 cuuuuuuuuu
 uuuuuuuuu
 uuuuuuuuu
 1800

 uuuuuuuuuu
 cuuucuugug
 gcuccaacuu
 agcecuagu
 caggauage
 gugaaagagu
 cgugagggg
 7800

 uuuuuuuuuuu
 cuuucuugggg
 cuccaacauu
 agcecuagu
 acggauage
 gugaaagggu
 cgugagecg
 7800

 uuucuugggg
 cuccaacau
 agcecuagu
 cucaagaag
 agguaaggg
 cgugaaggg
 7800

<210> 3

<211> 9678

<212> DNA

<213> Hepatitis C virus

<220>

<221> CDS

<222> (341)..(9442)

<400> 3

acetgeecet aataggggeg acacteegee atgaateact cecetgtgag gaactactgt 60

cttcacqcag aaagcgccta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120

cocceteccg ggagagecat agtggtetge ggaaceggtg agtacacegg aattgeeggg 180

aagactgggt cottlettgg ataaacccac totatgcocg gccatttggg cgtgcccccg 240

caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300

cgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat cct 355

Met Ser Thr Asn Pro

								1				5				
																402
			aga				-									403
Lys	Pro	Gln	Arg	Lys	Thr	Lys	Arg	Asn	Thr	Asn	Arg	Arg	Pro	Glu	Asp	
	10						15				20					
gtt	aag	tte	ccg	ggc	ggc	ggc	cag	atc	gtt	gge	gga	gta	tac	ttg	ttg	451
Va1	Lys	Phe	Pro	Gly	Gly	Gly	Gln	île	Val	Gly	Gly	Va1	Tyr	Leu	Leu	
	25					30	30 35									
ccg	cgc	agg	ggc	ccc	agg	ttg	ggt	gtg	cgc	acg	aca	agg	aaa	act	tog	499
Pro	Arg	Arg	Gly	Pro	Arg	Leu	Gly	Val	Arg	Thr	Thr	Arg	Lys	Thr	Ser	
		40				45				5	0					
gag	cgg	tee	cag	cca	cgt	ggg	aga	cgc	cag	ccc	atc	acc	aaa	gat	cgg	547
Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	Arg	Gln	Pro	Ile	Pro	Lys	Asp	Arg	
	55				60					55						
cgc	tee	act	ggc	aag	gec	tgg	gga	aaa	cca	ggt	aga	ccc	tgg	ccc	cta	595
Arg	Ser	Thr	Gly	ьув	Ala	Trp	Gly	Lys	Pro	Gly	Arg	Pro	Trp	Pro	Leu	
70				7	5				80				85			
tat	999	aat	gag	gga	ctc	ggc	tgg	gca	gga	tgg	ctc	ctg	taa	ccc	cga	643
Tyr	Gly	Asn	Glu	Gly	Leu	Gly	Trp	Ala	Gly	Trp	Leu	Leu	Ser	Pro	Arg	
			9	0				95				100				
ggc	tet	cgc	ccc	tcc	tgg	ggc	ccc	act	gac	ccc	cgg	cat	agg	teg	cgc	691
Gly	Ser	Arg	Pro	Ser	Trp	Gly	Pro	Thr	Asp	Pro	Arg	нів	Arg	Ser	Arg	
	105					110				1.1	5					

aac gtg ggt aaa gtc atc gac acc cta acg tgt ggc ttt gcc gac ctc 739 Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu

	:	130				125				13	0					
atg	999	tac	ate	ccc	gtc	gta	ggc	gcc	ccg	ctt	agt	ggc	gcc	gcc	aga	787
Met.	Gly	Tyr	Ile	Pro	Val	Val	Gly	Ala	Pro	Leu	Ser	Gly	Ala	Ala	Arg	
	135				140	0			1	45						
gct	gte	gcg	cac	ggc	gt g	aga	gtc	ctg	gag	gac	3 33	gtt	aat	tat	gca	835
Ala	Va 1	Ala	His	Gly	Val	Arg	Val	Leu	Glu	Asp	Gly	Va1	Asn	туг	Ala	
150				15	5			:	160				165			
aca	ggg	aac	cta	ccc	ggt	ttc	ccc	ttt	tet	atc	ttc	ttg	ctg	gcc	ctg	883
Thr	Gly	Asn	Leu	Pro	Gly	Phe	Pro	Phe	Ser	11e	Phe	Leu	Leu	Ala	Leu	
			17	0				175				180				
ttg	tee	tgc	atc	acc	gtt	ccg	gtc	tat	gct	gcc	cag	gtg	aag	aat	acc	931
Leu	Ser	Cys	Ile	Thr	Val	Pro	Va1	Ser	Ala	Ala	Gln	Va1	ьуз	Asn	Thr	
		1	85				190				19	5				
		1	85				190				19	5				
agt	agc			atg	gtg	acc		gac	tgc	tec			agc	atc	act	979
		agc	tac			acc Thr	aat				aat	gac				979
	Ser	agc	tac				aat				aat Asn	gac				979
	Ser	ago Ser	tac			Thr	aat			ser	aat Asn	gac				979
Ser	Ser	agc Ser	tac Tyr	Met	Val	Thr	aat Asn	Asp	Cys	Ser 21	aat Asn	gac Asp	Ser	Ile	Thr	979
Ser tgg	Ser 2 cag	ago Ser 800	tac Tyr	Met gct	Val gcg	Thr 205	aat Asn	Asp	Cys gtc	Ser 21 ccc	aat Asn O	gac Asp tgc	ser gtc	Ile	Thr	
Ser tgg Trp	Ser 2 cag	ago Ser 800	tac Tyr	Met gct	Val gcg	Thr 205 gtt Val	aat Asn	Asp	Cys gtc Val	Ser 21 ccc	aat Asn O	gac Asp tgc	ser gtc	Ile	Thr	
Ser tgg Trp	Ser cag Gln	ago Ser 800	tac Tyr	Met gct	Val gcg Ala	Thr 205 gtt Val	aat Asn	Asp	Cys gtc Val	Ser 21 ccc Pro	aat Asn O	gac Asp tgc	ser gtc	Ile	Thr	
ser tgg Trp	Ser cag Gln 215	agc Ser 200 ctc Leu	tac Tyr gag Glu	Met gct Ala	yal gcg Ala 220	Thr 205 gtt Val	aat Asn ctc Leu	Asp cac His	Cys gtc Val 2	ger 21 ccc Pro 25	aat Asn 0 ggg Gly	gac Asp tgc Cys	gtc Val	ccg Pro	Thr tgc Cys	
tgg Trp	cag Gln 215	agc Ser 200 ctc Leu	tac Tyr gag Glu	Met gct Ala	gcg Ala 220	Thr 205 gtt Val	aat Asn ctc Leu	cac His	gtc Val 2	ser 21 ccc Pro 25	aat Asn 0 ggg Gly	gac Asp tgc Cys	gtc Val	ccg Pro	Thr tgc Cys	1027
tgg Trp	cag Gln 215	agc Ser 200 ctc Leu	tac Tyr gag Glu	Met gct Ala	gcg Ala 220 acg	Thr 205 gtt Val	aat Asn ctc Leu	cac His tgt Cys	gtc Val 2	ser 21 ccc Pro 25	aat Asn 0 ggg Gly	gac Asp tgc Cys	gtc Val	ccg Pro	Thr tgc Cys	1027
tgg Trp gag Glu	cag Gln 215	agc Ser 200 ctc Leu	tac Tyr gag Glu	gct Ala aat Asn	gcg Ala 220 acg	Thr 205 gtt Val	aat Asn ctc Leu	cac His tgt Cys	gtc Val 2 tgg	ser 21 ccc Pro 25	aat Asn 0 ggg Gly	gac Asp tgc Cys	gtc Val tcg Ser	ccg Pro	Thr tgc Cys	1027
tgg Trp gag Glu 230	Cag Gln 215 aga Arg	agc Ser 200 ctc Leu gtg Val	tac Tyr gag Glu ggg	Met gct Ala aat Asn 23	yal gcg Ala 220 acg Thr	Thr 205 gtt Val	aat Asn ctc Leu cgg	cac His tgt Cys	gtc Val 2 tgg	Ser 23 ccc Pro 25 gtg Val	aat Asn 0 ggg Gly cca	gac Asp tgc Cys gtc Val	gtc Val tcg Ser 245	Ile ccg Pro	tgc Cys aac Asn	1027

			25	0			:	255				260				
atc	gat	atg	gtt	gtg	atg	tcc	gcc	acc	ttc	tgc	tet	gct	ctc	tac	gtg	1171
Ile	Asp	Met	Va1	Val	Met	Ser	Ala	Thr	Phe	Cys	Ser	Ala	Leu	Tyr	Val	
		2	65				270				27	5				
ggg	gac	ctc	tgt	ggc	ggg	gtg	atg	ctc	gcg	gcc	cag	gtg	ttc	atc	gtc	1219
Gly	Asp	Leu	Сув	Gly	Gly	Val	Met	Leu	Ala	Ala	Gln	Val	Phe	Ile	Val	
	2	8 0				285				29	0					
tcg	ccg	cag	tac	cac	tgg	ttt	gtg	caa	gaa	tgc	aat	tgc	tcc	atc	tac	1267
Ser	Pro	Gln	Tyr	His	Trp	Phe	Val	Gln	Glu	Сув	Asn	Сув	ser	Ile	Tyr	
	295				300)			3	05						
cct	ggc	acc	atc	act	gga	cac	ege	atg	gca	tgg	gac	atg	atg	atg	aac	1315
Pro	Gly	Thr	Ile	Thr	Gly	His	Arg	Met	Ala	Trp	Asp	Met	Met	Met	Asn	
310				31	5			:	320				325			
310				31	5			:	320				325			
	tcg	ccc	acg			atg	atc			tac	gtg	atg			ccc	1363
tgg			acg Thr	gcc	acc			ctg	geg				cgc	gte		1363
tgg				gcc Ala	acc		Ile	ctg	geg				cgc Arg	gte		1363
tgg			Thr	gcc Ala	acc		Ile	ctg Leu	geg			Met	cgc Arg	gte		1363
tgg Trp	Ser	Pro	Thr	gcc Ala 0	acc Thr	Met	Ile	ctg Leu 335	gcg Ala	Tyr	Val	Met 340	cgc Arg	gtc Val	Pro	1363
tgg Trp gag	Ser gtc	Pro	Thr 33	gcc Ala 0	acc Thr	Met gtt	Ile	ctg Leu 335	gcg Ala gct	Tyr	Val tgg	Met 340 ggc	cgc Arg	gtc Val	Pro	
tgg Trp gag	Ser gtc	Pro atc	Thr 33 ata	gcc Ala 0	acc Thr	Met gtt	Ile	ctg Leu 335	gcg Ala gct	Tyr	Val tgg	Met 340 ggc Gly	cgc Arg	gtc Val	Pro	
tgg Trp gag	Ser gtc	Pro atc	Thr 33 ata Ile	gcc Ala 0	acc Thr	Met gtt	lle agc Ser	ctg Leu 335	gcg Ala gct	Tyr	Val tgg Trp	Met 340 ggc Gly	cgc Arg	gtc Val	Pro	
tgg Trp gag Glu	ser gtc Val	Pro atc Ile	Thr 33 ata 11e 45	gcc Ala 0 gac Asp	acc Thr atc	Met gtt Val	agc Ser	ctg Leu 335 ggg Gly	gcg Ala gct Ala	Tyr cac His	tgg Trp 35	Met 340 ggc Gly 5	ege Arg gte Val	gtc Val atg Met	Pro	1411
tgg Trp gag Glu	ser gtc Val	Pro atc Ile 3	Thr 33 ata 11e 45	gcc Ala 0 gac Asp	acc Thr atc Ile	gtt Val	agc Ser 350	ctg Leu 335 ggg Gly	gcg Ala gct Ala	Tyr cac His	tgg Trp 35	Met 340 ggc Gly 5	cgc Arg gtc Val	gtc Val atg Met	Pro ttc Phe	1411
tgg Trp gag Glu	gtc Val ttg	Pro atc Ile 3	Thr 33 ata 11e 45	gcc Ala 0 gac Asp	acc Thr atc Ile	gtt Val	agc Ser 350	ctg Leu 335 ggg Gly	gcg Ala gct Ala	Tyr cac His	tgg Trp 35 gcg	Met 340 ggc Gly 5	cgc Arg gtc Val	gtc Val atg Met	Pro ttc Phe	1411
tgg Trp gag Glu	gtc Val ttg	Pro atc Ile 3 gcc Ala	Thr 33 ata 11e 45	gcc Ala 0 gac Asp	acc Thr atc Ile	Met Val atg	agc Ser 350	ctg Leu 335 ggg Gly	gcg Ala gct Ala	Tyr cac His tgg	tgg Trp 35 gcg	Met 340 ggc Gly 5	cgc Arg gtc Val	gtc Val atg Met	Pro ttc Phe	1411
tgg Trp gag Glu ggc	gtc Val ttg Leu	Pro atc Ile 3 gcc Ala	Thr 333 ata Ile 45 tac	gcc Ala 0 gac Asp	acc Thr atc Ile	gtt Val atg Met	agc Ser 350 cag	ctg Leu 335 ggg Gly gga Gly	gcg Ala gct Ala gcg Ala	Tyr cac His tgg Trp	tgg Trp 35 gcg Ala	Met 340 ggc Gly 5 aag	gtc Val	gtc Val atg Met	Pro ttc Phe	1411

	375				381	,			3	0.5							
															cat	1555	
	Ala	Val	Ala			Thr	Asn			Ala	Gly	Val			His		
390				39	5			4	100				405				
ggc	cot	cag	cag	aac	att	cag	ctc	att	aac	acc	aac	ggc	agt	tgg	cac	1603	
Gly	Pro	Gln			Ile	Gln			Asn	Thr	Asn			Trp	His		
			41	.0				115				420					
															ttt	1651	
Ile	Asn	Arg	Thr	Ala	Leu	Asn	Сув	Asn	Asp	Ser			Thr	Gly	Phe		
		4	25				430				4.3	5					
ctc	gcg	gcc	ttg	ttc	tac	acc	aac	cgc	ttt	aac	tcg	tca	999	tgt	cca	1699	
Leu	Ala	Ala	Leu	Phe	Tyr	Thr	Asn	Arg	Phe	Asn	ser	ser	Gly	Cys	Pro		
	4	4 0				445				4 5	0						
999	cgc	ctg	tcc	gcc	tgc	cgc	aac	atc	gag	gct	ttc	cgg	ata	999	tgg	1747	
Gly	Arg	Leu	Ser	Ala	Cys	Arg	Asn	Ile	Glu	Ala	Phe	Arg	Ile	Gly	Trp		
	455				460)			4	65							
ggc	acc	cta	cag	tac	gag	gat	aat	gtc	acc	aat	cca	gag	gat	atg	agg	1795	
Gly	Thr	Leu	Gln	Tyr	Glu	Asp	Asn	Val	Thr	Asn	Pro	Glu	Asp	Met	Arg		
470				47	5				180				485				
ccg	tac	tgc	tgg	cac	tac	ccc	cca	aag	ccg	tgt	ggc	gta	gt¢	ccc	gcg	1843	
Pro	Tyr	Сув	Trp	His	Tyr	Pro	Pro	Lys	Pro	Cys	Gly	Val	Val	Pro	Ala		
			49	0				195				500					
agg	tet	gtg	tgt	ggc	cca	gtg	tac	tgt	ttc	acc	ccc	agc	ccg	gta	gta	1891	

Arg Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser Pro Val Val

		5 (5				510				515	5				
gtg	ggc	acg	acc	gac	aga	cgt	gga	gtg	ccc	acc	tac	aca	tgg	gga	gag	1939
Val	Gly	Thr	Thr	Asp	Arg	Arg	Gly	Val	Pro	Thr	Tyr	Thr	Trp	Gly	Glu	
	5	20				525				53	0					
aat	gag	aca	gat	gtc	ttc	cta	ctg	aac	agc	acc	cga	ccg	ccg	cag	ggc	1987
Asn	Glu	Thr	Asp	Val	Phe	Leu	Leu	Asn	Ser	Thr	Arg	Pro	Pro	Gln	Gly	
	535				540	,			5	45						
tca	tgg	ttc	ggc	tgc	acg	tgg	atg	aac	tcc	act	ggt	ttc	acc	aag	act	2035
ser	Trp	Phe	Gly	Cys	Thr	Trp	Met	Asn	Ser	Thr	Gly	Phe	Thr	Lys	Thr	
550				55	5			5	60				565			
tgt	ggc	gcg	cca	cct	tgc	cgc	acc	aga	gct	gac	ttc	aac	gcc	agc	acg	2083
Сув	Gly	Ala	Pro	Pro	Cys	Arg	Thr	Arg	Ala	Asp	Phe	Asn	Ala	Ser	Thr	
			57	0				575				580				
gac	ttg	ttg	tgc	cct	acg	gat	tgt	ttt	agg	aag	cat	cct	gat	gcc	act	2131
Asp	Leu	Leu	Cys	Pro	Thr	Asp	Сув	Phe	Arg	Ьув	His	Pro	Asp	Ala	Thr	
		5	85				590				59	5				
tat	att	aag	tgt	ggt	tot	999	ccc	tgg	ctc	aca	cca	aag	tgc	ctg	gtc	2179
Tyr	Ile	Lys	Cys	Gly	Ser	Gly	Pro	Trp	Leu	Thr	Pro	Lys	Cys	Leu	Val	
	6	500				605				61	. 0					
cac	tac	cct	tac	aga	ctc	tgg	cat	tac	ccc	tgc	aca	gtc	aat	ttt	acc	2227
His	туr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Cys	Thr	v al	Asn	Phe	Thr	
	615				620)			6	25						
atc	ttc	aag	ata	aga	atg	tat	gta	aaa	999	gtt	gag	cac	agg	ctc	acg	2275

Ile Phe Lys Ile Arg Met Tyr Val Gly Gly Val Glu His Arg Leu Thr

630				63	5			6	40				645			
gcc	gca	tgc	aac	ttc	act	cgt	ggg	gat	cgc	tgc	gac	ttg	gag	gac	agg	2323
Ala	A1a	Сув	Asn	Phe	Thr	Arg	G1 y	Asp	Arg	Сув	Asp	Leu	Glu	Asp	Arg	
			65	0				655				660				
gac	agg	agt	cag	ctg	tct	cct	ctg	ttg	cac	tet	acc	acg	gaa	tgg	gac	2371
Asp	Arg	Ser	G1n	Leu	Ser	Pro	Leu	Leu	His	Ser	Thr	Thr	Glu	Trp	Ala	
		6	65				670				67	5				
atc	ctg	ccc	tgc	acc	tac	tca	gac	tta	ccc	gct	ttg	tca	act	ggt	ctt	2419
11e	Leu	Pro	Сув	Thr	Tyr	Ser	Авр	Leu	Pro	Ala	Leu	Ser	Thr	Gly	Leu	
	6	80				685				69	0					
ctc	cac	ctt	cac	cag	aac	atc	gtg	gac	gta	caa	tac	atg	tat	ggc	ctc	2467
Leu	His	Leu	His	Gln	Asn	Ile	Val	Asp	Val	Gln	Tyr	Met	Tyr	Gly	Leu	
	595				700)			7	05						
	595				700)			7	05						
		gct	atc	aca			gtc	gtt			gag	tgg	gtg	gta	ata	2515
tca	cct		atc		aaa	tac			cga	tgg						2515
tca	cct				aaa Lys	tac		Val	cga	tgg				Val		2515
tca Ser	cct			Thr	aaa Lys	tac		Val	cga Arg	tgg			Va1	Val		2515
tca Ser	cct Pro	Ala	Ile	Thr	aaa Lys 5	tac Tyr	Va1	Val	cga Arg 720	tgg Trp	G1 u	Trp	Val 725	Val		2515
tca Ser 710	cct Pro	Ala	Ile	Thr 71 tta	aaa Lys 5	tac Tyr	Val	Val aga	cga Arg 720 gtc	tgg Trp tgc	Glu gcc	Trp	Val 725 ttg	Val	Leu	
tca Ser 710	cct Pro	Ala	Ile	Thr 71 tta Leu	aaa Lys 5	tac Tyr	Val gcc Ala	Val aga	cga Arg 720 gtc	tgg Trp tgc	Glu gcc	Trp	Val 725 ttg Leu	Val	Leu	
tca Ser 710	cct Pro	Ala	Ile ctc Leu	Thr 71 tta Leu	aaa Lys 5	tac Tyr	Val gcc Ala	Val aga Arg	cga Arg 720 gtc	tgg Trp tgc	Glu gcc	Trp tgc Cys	Val 725 ttg Leu	Val	Leu	
tca Ser 710 tta Leu	cct Pro ttc Phe	Ala ctg Leu	ttc Leu 73	Thr 71 tta Leu 0	aaa Lys 5 gcg Ala	tac Tyr gac Asp	Val gcc Ala	Val aga Arg 735	cga Arg 720 gtc Val	tgg Trp tgc Cys	gcc Ala	tgc Cys	Val 725 ttg Leu	tgg Trp	Leu	
tca ser 710 tta Leu	cct Pro ttc Phe	Ala ctg Leu	ttc Leu 73	Thr 71 tta Leu 0	aaa Lys 5 gcg Ala	tac Tyr gac Asp	yal gcc Ala	aga Arg 735	cga Arg 720 gtc Val	tgg Trp tgc Cys	gcc Ala	tgc Cys 740	Val 725 ttg Leu	tgg Trp	Leu atg Met	2563
tca ser 710 tta Leu	cct Pro ttc Phe	Ala ctg Leu ttg	Ile ctc Leu 73	Thr 71 tta Leu 0	aaa Lys 5 gcg Ala	tac Tyr gac Asp	yal gcc Ala	aga Arg 735	cga Arg 720 gtc Val	tgg Trp tgc Cys	gcc Ala	tgc Cys 740 aag	Val 725 ttg Leu	tgg Trp	Leu atg Met	2563
tca ser 710 tta Leu	cct Pro ttc Phe	Ala ctg Leu ttg	ctc Leu 73 ttg	Thr 71 tta Leu 0	aaa Lys 5 gcg Ala	tac Tyr gac Asp	yal gcc Ala gaa Glu	aga Arg 735	cga Arg 720 gtc Val	tgg Trp tgc Cys	gcc Ala gag Glu	tgc Cys 740 aag	Val 725 ttg Leu	tgg Trp	Leu atg Met	2563
tca Ser 710 tta Leu ctc	cct Pro ttc Phe atc	Ala ctg Leu ttg Leu 7	Ile ctc Leu 73 ttg Leu	Thr 71 tta Leu 0 ggc Gly	aaa Lys 5 gcg Ala cag	tac Tyr gac Asp gcc	yal gcc Ala gaa Glu 750	val aga Arg 735 gca Ala	cga Arg 720 gtc Val	tgg Trp tgc Cys	gcc Ala gag Glu 75	tgc Cys 740 aag Lys	Val 725 ttg Leu ttg	tgg Trp	Leu atg Met	2563

760			765				77	0					
atc ttc ttc	gtg gca	gct	tgg	cac	atc	agg	ggt	cgg	gtg	gtc	ccc	ttg	2707
Ile Phe Phe	Val Ala	Ala	Trp	His	Ile	Arg	Gly	Arg	Val	Va1	Pro	Leu	
775		780	•			7	85						
acc acc tat	tgc ctc	act	ggc	cta	tgg	ccc	ttc	tgc	cta	ctg	ctc	atg	2755
Thr Thr Tyr	Cys Leu	Thr	Gly	Leu	Trp	Pro	Phe	Сув	Leu	Leu	Leu	Met	
790	79	5			8	300				805			
gca ctg ccc	egg cag	gct	tat	gcc	tat	gac	gca	cct	gtg	cac	gga	cag	2803
Ala Leu Pro	Arg Gln	Ala	Tyr	Ala	Tyr	Asp	Ala	Pro	Va 1	His	Gly	Gln	
	810				815				820				
ata ggc gtg	ggt ttg	ttg	ata	ttg	atc	acc	ctc	ttc	aca	ctc	acc	ccg	2851
Ile Gly Val	Gly Leu	Leu	Ile	Leu	Ile	Thr	Leu	Phe	Thr	Leu	Thr	Pro	
82	25			830				83	5				
ggg tat aag	acc ctc			cag				tgg	ttg				2899
ggg tat aag Gly Tyr Lys	acc ctc		Gly	cag			тгр	tgg Trp	ttg				2899
ggg tat aag	acc ctc			cag				tgg Trp	ttg				2899
ggg tat aag Gly Tyr Lys 840	acc ctc	Leu	Gly 845	cag Gln	Cys	ьeu	Trp	tgg Trp 0	ttg Leu	Сув	туг	Leu	
ggg tat aag Gly Tyr Lys 840 ctg acc ctg	acc ctc Thr Leu ggg gaa	Leu	Gly 845 atg	cag Gln att	Cys	Leu gag	Trp 85	tgg Trp 0	ttg Leu cca	Cys	Tyr	Leu	2899
ggg tat aag Gly Tyr Lys 840 ctg acc ctg Leu Thr Leu	acc ctc Thr Leu ggg gaa	Leu gcc Ala	Gly 845 atg Met	cag Gln att	Cys	Leu gag Glu	Trp 85 tgg Trp	tgg Trp 0	ttg Leu cca	Cys	Tyr	Leu	
ggg tat aag Gly Tyr Lys 840 ctg acc ctg	acc ctc Thr Leu ggg gaa	Leu	Gly 845 atg Met	cag Gln att	Cys	Leu gag Glu	Trp 85	tgg Trp 0	ttg Leu cca	Cys	Tyr	Leu	
ggg tat aag Gly Tyr Lys 840 ctg acc ctg Leu Thr Leu 855	acc ctc Thr Leu ggg gaa Gly Glu	gcc Ala 860	Gly 845 atg Met	cag Gln att Ile	Cys cag Gln	beu gag Glu 8	Trp 85 tgg Trp 65	tgg Trp 0 gta Val	ttg Leu cca Pro	ccc Pro	Tyr atg Met	Leu cag Gln	2947
ggg tat aag Gly Tyr Lys 840 ctg acc ctg Leu Thr Leu 855	acc ctc Thr Leu ggg gaa Gly Glu	gcc Ala 860 gat	Gly 845 atg Met	cag Gln att Ile	cag Gln gcg	gag Glu 8	tgg Trp 65	tgg Trp 0 gta Val	ttg Leu cca Pro	ccc Pro	Tyr atg Met	cag Gln tgc	
ggg tat aag Gly Tyr Lys 840 ctg acc ctg Leu Thr Leu 855 gtg cgc ggc Val Arg Gly	acc ctc Thr Leu ggg gaa Gly Glu ggc cgc Gly Arg	gcc Ala 860 gat Asp	Gly 845 atg Met	cag Gln att Ile	cag Gln gcg Ala	gag Glu 8 tgg	tgg Trp 65	tgg Trp 0 gta Val	ttg Leu cca Pro	ccc Pro	atg Met ttc	cag Gln tgc	2947
ggg tat aag Gly Tyr Lys 840 ctg acc ctg Leu Thr Leu 855	acc ctc Thr Leu ggg gaa Gly Glu	gcc Ala 860 gat Asp	Gly 845 atg Met	cag Gln att Ile	cag Gln gcg Ala	gag Glu 8	tgg Trp 65	tgg Trp 0 gta Val	ttg Leu cca Pro	ccc Pro	atg Met ttc	cag Gln tgc	2947
ggg tat aag Gly Tyr Lys 840 ctg acc ctg Leu Thr Leu 855 gtg cgc ggc Val Arg Gly 870	acc ctc Thr Leu ggg gaa Gly Glu ggc cgc Gly Arg 87	gcc Ala 860 gat Asp	Gly 845 atg Met	cag Gln att Ile atc	cag Gln gcg Ala	gag Glu 8 tgg Trp	tgg trp 65	tgg Trp 0 gta Val	ttg Leu cca Pro	ccc Pro	Tyr atg Met ttc	cag Gln tgc Cys	2947 2995
ggg tat aag Gly Tyr Lys 840 ctg acc ctg Leu Thr Leu 855 gtg cgc ggc Val Arg Gly	acc ctc Thr Leu ggg gaa Gly Glu ggc cgc Gly Arg 87	gcc Ala 8600 gat Asp 5	Gly 845 atg Met	cag Gln att Ile atc Ile	cag Gln gcg Ala {	gag Glu 8 tgg Trp	tgg trp 65 gcc Ala	tgg Trp 0 gta Val gtc Val	ttg Leu cca Pro act Thr	ccc Pro	Tyr atg Met ttc Phe	cag Gln tgc Cys	2947

	890	895	900	
cct gct tac c	to tta agg goo	gct ttg aca	cat gtg ccg tac	ttc gtc 3091
Pro Ala Tyr Le	eu Leu Arg Ala	Ala Leu Thr	His Val Pro Tyr	Phe Val
905		910	915	
aga get cac g	ct ctg ata ago	gta tgc gct	ttg gtg aag cag	ctc gcg 3139
Arg Ala His A	la Leu Ile Arç	Val Cys Ala	Leu Val Lys Glr	Leu Ala
920	92	5	930	
ggg ggt agg ta	at gtt cag gtg	geg eta ttg	gee ett gge agg	tgg act 3187
Gly Gly Arg T	yr Val Gln Val	. Ala Leu Leu	Ala Leu Gly Arg	Trp Thr
935	940	9	4.5	
			atg teg gac tgg	
Gly Thr Tyr I	le Tyr Asp His	Leu Thr Pro	Met Ser Asp Trp	Ala Ala
950	955	960	96	5
ago ggo otg og	gc gac tta gcg	gte gee gtg	gaa ccc atc atc	ttc agt 3283
ago ggo otg og	go gao tta gog rg Asp Leu Ala	gtc gcc gtg	gaa ccc atc atc	ttc agt 3283
ago ggo otg og	gc gac tta gcg	gte gee gtg	gaa ccc atc atc	ttc agt 3283
age gge etg eg Ser Gly Leu A	gc gac tta gcg rg Asp Leu Ala 970	gtc gcc gtg Val Ala Val 975	gaa ccc atc atc Glu Pro Ile Ile 980	ttc agt 3283 Phe Ser
agc ggc ctg co	gc gac tta gcg rg Asp Leu Ala 970 ag aag gtc atc	gtc gcc gtg Val Ala Val 975	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct	ttc agt 3283 Phe Ser gca tgt 3331
age gge etg eg Ser Gly Leu A: ceg atg gag at	gc gac tta gcg rg Asp Leu Ala 970 ag aag gtc atc	gtc gcc gtg Val Ala Val 975 gtc tgg gga	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct Ala Glu Thr Ala	ttc agt 3283 Phe Ser gca tgt 3331
agc ggc ctg co	gc gac tta gcg rg Asp Leu Ala 970 ag aag gtc atc	gtc gcc gtg Val Ala Val 975	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct	ttc agt 3283 Phe Ser gca tgt 3331
agc ggc ctg cg Ser Gly Leu As ccg atg gag ac Pro Met Glu Ly 985	gc gac tta gcg rg Asp Leu Als 970 ag aag gtc atc ys Lys Val Ile	gtc gcc gtg Val Ala Val 975 gtc tgg gga Val Trp Gly 990	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct Ala Glu Thr Ala 995	ttc agt 3283 Phe Ser gca tgt 3331 Ala Cys
age gge etg eg Ser Gly Leu Ar cog atg gag an Pro Met Glu Ly 985	gc gac tta gog rg Asp Leu Als 970 ag aag gtc atc ys Lys Val Ile ta cat gga ctt	gtc gcc gtg Val Ala Val 975 gtc tgg gga Val Trp Gly 990	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct Ala Glu Thr Ala 995	ttc agt 3283 Phe Ser gca tgt 3331 Ala Cys cag gag 3379
agc ggc ctg cq Ser Gly Leu A: ccg atg gag aa Pro Met Glu Ly 985 ggg gac att ct Gly Asp Ile Le	gc gac tta gcg rg Asp Leu Als 970 ag aag gtc atc ys Lys Val Ile ta cat gga ctt eu His Gly Leu	gtc gcc gtg Val Ala Val 975 gtc tgg gga Val Trp Gly 990 ccc gtg tcc	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct Ala Glu Thr Ala 995 gcc cga ctc ggc	ttc agt 3283 Phe Ser gca tgt 3331 Ala Cys cag gag 3379
age gge etg eg Ser Gly Leu Ar cog atg gag an Pro Met Glu Ly 985	gc gac tta gog rg Asp Leu Als 970 ag aag gtc atc ys Lys Val Ile ta cat gga ctt	gtc gcc gtg Val Ala Val 975 gtc tgg gga Val Trp Gly 990 ccc gtg tcc	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct Ala Glu Thr Ala 995	ttc agt 3283 Phe Ser gca tgt 3331 Ala Cys cag gag 3379
agc ggc ctg c; Ser Gly Leu A: ccg atg gag at Pro Met Glu L; 985 ggg gac att ct Gly Asp Ile Le	gc gac tta gcg rg Asp Leu Als 970 ag aag gtc atc ys Lys Val Ile ta cat gga ctt cu His Gly Leu 100	gtc gcc gtg Val Ala Val 975 gtc tgg gga Val Trp Gly 990 ccc gtg tcc Pro Val Ser	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct Ala Glu Thr Ala 995 gcc cga ctc ggc Ala Arg Leu Gly 1010	ttc agt 3283 Phe Ser goa tgt 3331 Ala Cys cag gag 3379 Gln Glu
age gge etg eg Ser Gly Leu A: ceg atg gag at Pro Met Glu L; 985 ggg gac att et Gly Amp Ile Le 1000 ate ete ete eg	gc gac tta gcg rg Asp Leu Als 970 ag aag gtc atc rg Lys Val Ile ta cat gga ctt cu His Gly Leu 100	gtc gcc gtg Val Ala Val 975 gtc tgg gga Val Trp Gly 990 ccc gtg tcc Pro Val Ser 5	gaa ccc atc atc Glu Pro Ile Ile 980 gcg gag acg gct Ala Glu Thr Ala 995 gcc cga ctc ggc	ttc agt 3283 Phe Ser 3331 gca tgt 3331 Ala Cys cag gag 3379 Gln Glu 3427

1015	1020	1	025	
ctt gct ccc atc	not got tot	200 CB0 CB3	aca caa aac	ctc ctq qqc 3475
Leu Ala Pro Ile				
1030	1035	1040		1045
gcc ata gtg gtg	agt atg acg	ggg cgt gac	agg aca gaa	cag gcc ggg 3523
Ala Ile Val Val	Ser Met Thr	Gly Arg Asp		
105	0	1055	1060)
gaa gtc caa atc	aha haa 202	ata tat asa	taa tta ata	gga aca acc 3571
Glu Val Gln Ile				33
1065		1070	1075	•
atc tcg ggg gtt	ttg tgg act	gtt tac cac	gga get gge	aac aag act 3619
Ile Ser Gly Val	Leu Trp Thr	Val Tyr His	Gly Ala Gly	Asn Lys Thr
		_		
1080	108	5	1090	
cta gcc ggc tta	cgg ggt ccg	gtc acg cag	atg tac tcg	
cta gcc ggc tta Leu Ala Gly Leu	cgg ggt ccg	gtc acg cag	atg tac tcg	-3- 3 3
cta gcc ggc tta	cgg ggt ccg	gtc acg cag	atg tac tcg	-3- 3 3
cta gcc ggc tta Leu Ala Gly Leu	egg ggt eeg Arg Gly Pro	gtc acg cag Val Thr Gln	g atg tac teg Met Tyr Ser 105	Ser Ala Glu
cta gcc ggc tta Leu Ala Gly Leu 1095	cgg ggt ccg Arg Gly Pro 1100	gtc acg cag Val Thr Gln 1	s atg tac tcg Met Tyr Ser 105	Ser Ala Glu
cta gcc ggc tta Leu Ala Gly Leu 1095 ggg gac ttg gta	cgg ggt ccg Arg Gly Pro 1100	gtc acg cag Val Thr Gln 1	s atg tac tcg Met Tyr Ser 105	Ser Ala Glu
cta gcc ggc tta Leu Ala Gly Leu 1095 ggg gac ttg gta Gly Asp Leu Val	cgg ggt ccg Arg Gly Pro 1100 ggc tgg ccc Gly Trp Pro	g gtc acg cag val Thr Gln 1 agc ccc cct Ser Pro Pro	s atg tac tcg Met Tyr Ser 105 ggg acc aag Gly Thr Lys	Ser Ala Glu tot ttg gag 3715 Ser Leu Glu 1125
cta gcc ggc tta Leu Ala Gly Leu 1095 ggg gac ttg gta Gly Asp Leu Val 1110 ccg tgc aag tgt	cgg ggt ccg Arg Gly Pro 1100 ggc tgg ccc Gly Trp Pro 1115	g gtc acg cag y Val Thr Gln 1 c agc ccc cct Ser Pro Pro 1120	atg tac tcg Met Tyr Ser 105 ggg acc aag Gly Thr Lys	Ser Ala Glu tot ttg gag 3715 Ser Leu Glu 1125 cgg aac gct 3763
cta gcc ggc tta Leu Ala Gly Leu 1095 ggg gac ttg gta Gly Asp Leu Val 1110 ccg tgc aag tgt Fro Cys Lys Cys	cgg ggt ccg Arg Gly Prc 1100 ggc tgg ccc Gly Trp Prc 1115 gga gcc gtc Gly Ala Val	gtc acg cag val Thr Gln 1 c agc ccc ccc Ser Pro Pro 1120 c gac cta tat	g atg tac tcg Met Tyr Ser 105 ggg acc aag Gly Thr Lys ctg gtc acg	Ser Ala Glu tot ttg gag 3715 Ser Leu Glu 1125 cgg aac get 3763 Arg Asn Ala
cta gcc ggc tta Leu Ala Gly Leu 1095 ggg gac ttg gta Gly Asp Leu Val 1110 ccg tgc aag tgt	cgg ggt ccg Arg Gly Prc 1100 ggc tgg ccc Gly Trp Prc 1115 gga gcc gtc Gly Ala Val	g gtc acg cag y Val Thr Gln 1 c agc ccc cct Ser Pro Pro 1120	atg tac tcg Met Tyr Ser 105 ggg acc aag Gly Thr Lys	Ser Ala Glu tot ttg gag 3715 Ser Leu Glu 1125 cgg aac get 3763 Arg Asn Ala
cta gcc ggc tta Leu Ala Gly Leu 1095 ggg gac ttg gta Gly Asp Leu Val 1110 ccg tgc aag tgt Fro Cys Lys Cys	cgg ggt ccg Arg Gly Pro 1100 ggc tgg ccc Gly Trp Pro 1115 gga gcc gtc Cly Ala Val	gtc acg cag Val Thr Gln 1 c agc ccc cct 1120 c gac cta tat 1 Asp Leu Tyr	g atg tac tcg Met Tyr Ser 105 ggg acc aag Gly Thr Lys ctg gtc acg Leu Val Thr	Ser Ala Glu tot ttg gag 3715 Ser Leu Glu 1125 cgg aac got 3763 Arg Asn Ala

tcc ccg aga	ccc att tcg	acc ttg aag	ggg tcc tcg	ggg ggg ccg	gtg 3859
Ser Pro Arg	Pro Ile Ser	Thr Leu Lys	Gly Ser Ser	Gly Gly Pro	Val
1160		1165	1170		
ctc tgc cct	agg ggc cac	gtc gtt ggg	ctc ttc cga	gca gct gtg	tgc 3907
Leu Cys Pro	Arg Gly His	Val Val Gly	Leu Phe Arg	Ala Ala Val	Сув
1175	1180	0	1185		
tot ogg ggo	gtg gcc aaa	tee ate gat	ttc atc ccc	gtt gag aca	ctc 3955
Ser Arg Gly	Val Ala Lys	Ser Ile Asp	Phe Ile Pro	Val Glu Thr	Leu
1190	1195	:	1200	1205	
gac gtt gtt	aca agg tct	dec act tte	agt gac aac	age acg cca	ccg 4003
Asp Val Val	Thr Arg Ser	Pro Thr Phe	Ser Asp Asn	Ser Thr Pro	Pro
	1210	1215		1220	
get gtg ccc	cag acc tat	cag gtc ggg	tac ttg cat	get eea aet	ggc 4051
Ala Val Pro	Gln Thr Tyr	Gln Val Gly	Tyr Leu His	Ala Pro Thr	Gly
122	15	1230	123	5	
			gcg tat gcc		
	-		Ala Tyr Ala	Ala Gln Gly	Tyr
1240		1245	1250		
-			get gee ace		
•			Ala Ala Thr	Leu Gly Phe	Gly
1255	126	0	1265		
			aat ccc aac		
Ala Tyr Leu	Ser Lys Ala	His Gly Ile	Asn Pro Asn	Ile Arg Thr	Gly

1270	1275	;	1280	1285	
				tee aca tat gge 4	243
val Alg in	1290	1295		1300	
	1290	1295		1300	
ttt ata	and gat ga	+	200 000 000	tat gac atc atc 4	291
	-			Tyr Asp Ile Ile	
	. Ala Asp Gi	1310	1315		
1.	005	1310	1315	'	
			ant see tee	att ata aga ata 4	1339
					1333
	сти сув ит		1330	Ile Leu Gly Ile	
1320		1325	1330		
					1387
					1307
-				Arg Leu Thr Val	
1335	13	10	1345		
	-			-	1435
Leu Ala Thr	Ala Thr Pro	Pro Gly Ser	Val Thr Thr	Pro His Pro Asp	
1350	1355		1360	1365	
ata gaa gag	gta ggc ct	ggg cgg gag	ggt gag atc	occ tto tat ggg 4	1483
Ile Glu Glu	Val Gly Le	Gly Arg Glu	Gly Glu Ile	Pro Phe Tyr Gly	
	1370	1375		1380	
agg gcg att	ccc cta to	tgc atc aag	gga ggg aga	cac ctg att ttc 4	1531
Arg Ala Ile	Pro Leu Se	Cys Ile Lys	Gly Gly Arg	His Leu Ile Phe	
1:	885	1390	1395	•	
tgc cac tca	aag aaa aa	tgt gac gag	ctc gcg gcg	goe ett egg gge	1579
Cys His Ser	Lys Lys Ly	Cys Asp Glu	Leu Ala Ala	Ala Leu Arg Gly	

1400		1405			141	LO					
										- 4	4600
atg ggc ttg											4627
Met Gly Leu			TYT TY		25	neu	Asp	vaı	ser	116	
1415	1	1420		14	125						
ata cca gct		*** ***	ata at	a ata	~~~	200	~=~	acc	ata	ato	4675
Ile Pro Ala											
1430	1435		· · · · · · · · · · · · · · · · · · ·	1440	*****			144			
1430	1433			2440							
acg ggg tac	act gga g	ac ttt	gac to	c ata	atc	qac	tgc	aat	gta	gag	4723
Thr Gly Tyr											
	1450	-	145			-	146				
gtc acc caa	gct gtc g	gac ttc	agc ct	g gac	ccc	acc	ttc	act	ata	acc	4771
Val Thr Gln	Ala Val A	ap Phe	Ser Le	u Asp	Pro	Thr	Phe	Thr	Ile	Thr	
14	65	1	470			147	5				
14	65	1	470			147	5				
14				c tca	ege			cgc	cgc	a aa	4819
	gtc cca c	caa gac	gct gt			agt	cag				4819
aca cag act	gtc cca c	caa gac	gct gt			agt Ser	cag				4819
aca cag act	gtc cca c	caa gac 31n Asp	gct gt		Arg	agt Ser	cag				4819
aca cag act	gtc cca c	caa gac 31n Asp 1485	gct gt Ala Va	ıl Ser	Arg	agt Ser	cag Gln	Arg	Arg	Gly	4819
aca cag act Thr Gln Thr 1480	gtc cca c Val Pro C	caa gac 31n Asp 1485 aga cag	gct gt Ala Va ggc ac	al Ser	Arg 14:	agt Ser 90	cag Gln gtt	Arg	Arg	Gly ggt	
aca cag act Thr Gln Thr 1480	gtc cca c	caa gac 31n Asp 1485 aga cag	gct gt Ala Va ggc ac	al Ser	Arg 14:	agt Ser 90	cag Gln gtt	Arg	Arg	Gly ggt	
aca cag act Thr Gln Thr 1480 cgc aca ggt Arg Thr Gly	gtc cca c	caa gac 31n Asp 1485 aga cag Arg Gln	gct gt Ala Va ggc ac	al Ser	Arg 14: agg Arg	agt Ser 90	cag Gln gtt	Arg	Arg	Gly ggt	
aca cag act Thr Gln Thr 1480 cgc aca ggt Arg Thr Gly	gtc cca c	caa gac 31n Asp 1485 aga cag Arg Gln	gct gt Ala Va ggc ac Gly Th	t tat t tat Tyr	Arg 14: agg Arg 505	agt Ser 90 tat Tyr	cag Gln gtt Val	tcc Ser	Arg act Thr	ggt ggt	
aca cag act Thr Gln Thr 1480 cgc aca ggt Arg Thr Gly 1495	gtc cca c Val Pro C aga gga a Arg Gly A	raa gac 31n Asp 1485 aga cag Arg Gln 1500	get gt Ala Ve gge ac Gly Th	t tat or Tyr 19	Arg 14: agg Arg 505	agt Ser 90 tat Tyr	cag Gln gtt Val	tcc ser	act Thr	ggt Gly tac	4867
aca cag act Thr Gln Thr 1480 cgc aca ggt Arg Thr Gly 1495 gaa cga gcc	gtc cca c Val Pro C aga gga a Arg Gly A	taa gac 31n Asp 1485 aga cag arg Gln 1500	get gt Ala Ve gge ac Gly Th	t tat or Tyr 19	Arg 14: agg Arg 505	agt Ser 90 tat Tyr	cag Gln gtt Val	tcc ser	act Thr tgc Cys	ggt Gly tac	4867
aca cag act Thr Gln Thr 1480 cgc aca ggt Arg Thr Gly 1495 gaa cga gcc Glu Arg Ala 1510	gtc cca c Val Pro C aga gga a Arg Gly A tca gga a Ser Gly M	caa gac elln Asp 1485 aga cag arg Gln 1500 atg ttt det Phe	get gt Ala Ve ggc ac Gly Th gac ac	at tat ar Tyr 15 gt gta ar Val	agg Arg OS gtg Val	agt Ser 90 tat Tyr	cag Gln gtt Val tgt	tcc Ser gag Glu 152	act Thr tgc Cys	ggt Gly tac	4867
aca cag act Thr Gln Thr 1480 cgc aca ggt Arg Thr Gly 1495 gaa cga gcc Glu Arg Ala	gtc coa c Val Pro C aga gga a Arg Gly A tca gga a Ser Gly M 1515	caa gac 3ln Asp 1485 aga cag Arg Gln 1500 atg ttt 4et Phe	gct gt Ala Ve ggc ac Gly Th gac ac Asp Se	et tat er Tyr 19 gt gta er Val 1520	Arg 14: agg Arg 605 gtg Val	agt Ser 90 tat Tyr ctt Leu	cag Gln gtt Val tgt Cys	tcc Ser gag Glu 152	act Thr tgc Cys 5	ggt gly tac Tyr	4867

1530	1535	1540	
agg ctt aga gcg tat	tto aac acg ccc g	ge eta cee gtg tgt c	caa gac 5011
Arg Leu Arg Ala Tyr	Phe Asn Thr Pro G	ly Leu Pro Val Cys G	31n Asp
1545	1550	1555	
cat ctt gaa ttt tgg	gag gca gtt ttc a	co gge oto aca cac a	ita gac 5059
His Leu Glu Phe Trp	Glu Ala Val Phe T	hr Gly Leu Thr His I	Ile Asp
1560	1565	1570	
gcc cac ttc ctc tcc	caa aca aag caa g	cg ggg gag aac ttc g	gcg tac 5107
Ala His Phe Leu Ser	Gln Thr Lys Gln A	la Gly Glu Asn Phe A	Ala Tyr
1575	1580	1585	
cta gta gcc tac caa	get acg gtg tgc g	cc aga gcc aag gcc	oct ccc 5155
Leu Val Ala Tyr Gln	Ala Thr Val Cys A	la Arg Ala Lys Ala i	Pro Pro
1590 159	5 160	0 1605	
ceg tee tgg gae gee	atg tgg aag tgc c	tg gcc cga ctc aag o	oct acg 5203
ccg tcc tgg gac gcc Pro Ser Trp Asp Ala			
Pro Ser Trp Asp Ala	Met Trp Lys Cys L	eu Ala Arg Leu Lys 1	
Pro Ser Trp Asp Ala	Met Trp Lys Cys L	eu Ala Arg Leu Lys 1	Pro Thr
Pro Ser Trp Asp Ala 1610	Met Trp Lys Cys L 1615 cct ctc ctg tac c	eu Ala Arg Leu Lys 1 1620 gt ttg ggc cct att a	Pro Thr
Pro Ser Trp Asp Ala 1610 ctt geg ggc ccc aca	Met Trp Lys Cys L 1615 cct ctc ctg tac c	eu Ala Arg Leu Lys 1 1620 gt ttg ggc cct att a	Pro Thr
Pro Ser Trp Asp Ala 1610 ctt gcg ggc ccc aca Leu Ala Gly Pro Thr	Met Trp Lys Cys L 1615 cct ctc ctg tac c Pro Leu Leu Tyr A	eu Ala Arg Leu Lys 1 1620 gt ttg ggc cct att 2 rg Leu Gly Pro Ile 2	Pro Thr
Pro Ser Trp Asp Ala 1610 ctt gcg ggc ccc aca Leu Ala Gly Pro Thr	Met Trp Lys Cys L 1615 cct ctc ctg tac c Pro Leu Leu Tyr A 1630	eu Ala Arg Leu Lys 1 1620 gt ttg ggc cct att a rg Leu Gly Pro Ile 1 1635	Pro Thr acc aat 5251 Thr Asn
Pro Ser Trp Asp Ala 1610 ctt geg gge ccc aca Leu Ala Gly Pro Thr 1625	Met Trp Lys Cys L 1615 cct ctc ctg tac c Pro Leu Leu Tyr A 1630 cac cct ggg acg a	gt ttg ggc cct att a rg Leu Gly Pro Ile ? 1635	Pro Thr acc aat 5251 Thr Asn
Pro Ser Trp Asp Ala 1610 ctt geg gge ccc aca Leu Ala Gly Pro Thr 1625 gag gtc acc ctc aca	Met Trp Lys Cys L 1615 cct ctc ctg tac c Pro Leu Leu Tyr A 1630 cac cct ggg acg a	gt ttg ggc cct att a rg Leu Gly Pro Ile ? 1635	Pro Thr acc aat 5251 Thr Asn
Pro Ser Trp Asp Ala 1610 ctt gog gge ccc aca Leu Ala Gly Pro Thr 1625 gag gtc acc ctc aca Glu Val Thr Leu Thr	Met Trp Lys Cys L 1615 cct ctc ctg tac c Pro Leu Leu Tyr A 1630 cac cct ggg acg a His Pro Gly Thr L	gt ttg ggc cct att a rg Leu Gly Pro Ile ? 1635 ag tac atc gcc aca a ys Tyr Ile Ala Thr (Pro Thr acc aat 5251 Thr Asn
Pro Ser Trp Asp Ala 1610 ctt gog gge ccc aca Leu Ala Gly Pro Thr 1625 gag gtc acc ctc aca Glu Val Thr Leu Thr	Met Trp Lys Cys L 1615 cct ctc ctg tac c Pro Leu Leu Tyr A 1630 cac cct ggg acg a His Pro Gly Thr L	eu Ala Arg Leu Lys 1 1620 gt ttg ggc cct att a rg Leu Gly Pro Ile ? 1635 ag tac atc gcc aca a ye Tyr Ile Ala Thr 6	ero Thr acc aat 5251 Thr Asn tgc atg 5299

gtc ctg	gca	gcc	gte	gcc	gca	tat	tgc	ctg	gcg	act	gga	tgc	gtt	tcc	5395
Val Leu	Ala	Ala	Val	Ala	Ala	Tyr	Сув	Leu	Ala	Thr	Gly	Cys	Val	Ser	
1670			167	5			1	680				168	5		
atc atc	ggc	ege	ttg	cac	gtc	aac	cag	cga	gtc	gtc	gtt	gcg	ccg	gat	5443
Ile Ile	Gly	Arg	Leu	His	Val	Asn	Gln	Arg	Va1	Val	Val	Ala	Pro	Asp	
		169	0			1	695				1700	0			
aag gag	gtc	ctg	tat	gag	gct	ttt	gat	gag	atg	gag	gaa	tgc	gcc	tct	5491
Lys Glu	Val	Leu	Tyr	G1u	Ala	Phe	Asp	Glu	Met	Glu	Glu	Сув	Ala	Ser	
	17	05			1	1710				171	5				
agg gcg	gct	ctc	atc	gaa	gag	ggg	cag	cgg	ata	gcc	gag	atg	ttg	aag	5539
Arg Ala	Ala	Leu	Ile	Glu	Glu	Gly	Gln	Arg	Ile	Ala	Glu	Met	Leu	Lys	
17	720				1725				17	30					
11	720				1725				17	30					
tcc aag		caa	gg¢				cag	gcc			cag	gec	cag	gac	5587
	atc			ttg	ctg	cag			tet	aag					5587
tec aag	atc			ttg	ctg Leu	cag		Ala	tet	aag					5587
tcc aag Ser Lys	atc			ttg Leu	ctg Leu	cag		Ala	tct Ser	aag					5587
tcc aag Ser Lys	atc Ile	Gln	Gly	ttg Leu 174	ctg Leu	cag Gln	Gln	Ala 17	tct Ser	aag Lys	Gln	Ala	Gln	Asp	5587
tec aag Ser Lys 1735	atc Ile	Gln gct	Gly atg	ttg Leu 174	ctg Leu 0	cag Gln tca	Gln tgg Trp	Ala 17 ccc Pro	tct ser 45	aag Lys gtg	Gln gaa	Ala caa Gln	Gln ttt Phe	Aap tgg	
tcc aag Ser Lys 1735 ata caa	atc Ile	Gln gct	Gly atg	ttg Leu 174 cag	ctg Leu 0	cag Gln tca	Gln tgg Trp	Ala 17	tct ser 45	aag Lys gtg	Gln gaa	Ala	Gln ttt Phe	Aap tgg	
tcc aag Ser Lys 1735 ata caa Ile Gln	atc Ile	Gln gct	Gly atg Met	ttg Leu 174 cag	ctg Leu 0	cag Gln tca	Gln tgg Trp	Ala 17 ccc Pro	tct ser 45	aag Lys gtg	Gln gaa	Ala caa Gln	Gln ttt Phe	Aap tgg	
tcc aag Ser Lys 1735 ata caa Tle Gln 1750 gcc aga	atc Ile ccc Pro	gct Ala	atg Met 175	ttg Leu 174 cag Gln 55	ctg Leu 0 gct Ala	cag Gln tca ser	tgg Trp 1	Ala 11 ccc Pro 760	tct Ser 45 aaa Lys	aag Lys gtg Val	Gln gaa Glu tac	caa Gln 176	Gln ttt Phe 5	tgg Trp	
tcc aag Ser Lys 1735 ata caa Ile Gln 1750	atc Ile ccc Pro	gct Ala atg	atg Met 175 tgg	ttg Leu 174 cag Gln 55	ctg Leu 0 gct Ala	cag Gln tca ser	tgg Trp 1 agc ser	Ala 11 ccc Pro 760	tct Ser 45 aaa Lys	aag Lys gtg Val	gaa Glu tac Tyr	caa Gln 176 ctc	Gln ttt Phe 5	tgg Trp	5635
tcc aag Ser Lys 1735 ata caa Tle Gln 1750 gcc aga	atc Ile ccc Pro	gct Ala	atg Met 175 tgg	ttg Leu 174 cag Gln 55	ctg Leu 0 gct Ala	cag Gln tca ser	tgg Trp 1	Ala 11 ccc Pro 760	tct Ser 45 aaa Lys	aag Lys gtg Val	Gln gaa Glu tac	caa Gln 176 ctc	Gln ttt Phe 5	tgg Trp	5635
tcc aag Ser Lys 1735 ata caa Tle Gln 1750 gcc aga Ala Arg	atc Ile ccc Pro cac	Gln gct Ala atg Met	atg Met 179 tgg Trp	Leu 174 cag Gln 55	ctg Leu 0 gct Ala	cag Gln tca ser att	tgg Trp 1 agc ser 775	Ala 17 ccc Pro 760 ggc Gly	tct Ser 45 aaa Lys atc	aag Lys gtg Val caa	gaa Glu tac Tyr	caa Gln 176 ctc Leu	ttt Phe 5	tgg Trp gga Gly	5635 5683
tcc aag Ser Lys 1735 ata caa Tle Gln 1750 gcc aga	atc Ile ccc Pro cac	gct Ala atg Met 1777	atg Met 179 tgg Trp 0	ttg Leu 174 cag Gln 55 aac Asn	ctg Leu 0 gct Ala ttc Phe	cag Gln tca Ser att Ile 1	tgg Trp 1 agc ser 775	Ala 11 ccc Pro 760 ggc Gly	tct Ser 45 aaa Lys atc Ile	aag Lys gtg Val caa Gln	gaa Glu tac Tyr 1786	caa Gln 176 ctc Leu	Gln tttt Phe S gca Ala	tgg Trp gga Gly	5635

agt	gcc	gcc	ctc	acc	agt	ccg	ttg	tcg	acc	agt	acc	acc	atc	ctt	ctc	5779
Ser	Ala	Ala	Leu	Thr	Ser	Pro	Leu	ser	Thr	Ser	Thr	Thr	Ile	Leu	Leu	
	18	300				1805				18	10					
aac	atc	atg	gga	ggc	tgg	tta	gcg	tac	cag	atc	gca	cca	ccc	gcg	ggg	5827
Asn	Ile	Met	Gly	Gly	Trp	Leu	Ala	Ser	Gln	Ile	Ala	Pro	Pro	Ala	Gly	
1	815				182	0			1.8	25						
qcc	acc	ggc	ttt	gtc	gtc	agt	ggc	ctg	gtg	999	gct	gcc	gtg	ggc	agc	5875
		Gly														
183				183			•		840				184			
ata	aac	cta	aat	220	ata	et a	ata	gac	ato	cha	gca	gga	tat	αat	g¢g	5923
		Leu														
116	GIY	Бец	185	-	*41	пец		855	110	Deu	ALU	186		G.J		
			100				•	055				100				
													~~~	~~~	220	5971
		_		-											aag	3371
GIA	тте	Ser	-	ATS	Leu			Pne	гуя	TIE			GIĀ	GIU	гув	
		18	65			-	1870				187	5				
		-	-	-	-				-				-		eeg	6019
Pro		Met	Glu	Asp	Val			Leu	Leu			Ile	Leu	Ser	Pro	
	18	8 0				1885				18	90					
gga	gcc	ctg	gtg	gtg	ggg	gtc	atc	tgc	gcg	gcc	att	ctg	cgc	cgc	cac	6067
Gly	Ala	Leu	Val	Va1	Gly	Val	Ile	Cys	Ala	Ala	Ile	Leu	Arg	Arg	His	
1	895				190	0			19	05						
gtg	gga	ccg	<b>3</b> 33	gag	ggc	geg	gtc	caa	tgg	atg	aac	agg	ctt	att	gcc	6115
Va1	Gly	Pro	Gly	Glu	Gly	Ala	Val	Gln	Trp	Met	Asn	Arg	Leu	Ile	Ala	

ttt	act	tee	aga	gga	aac	cac	atc	acc	cat	act	cac	tac	ata	асч	qaq	6163
			Arg													
			193					935				1940				
teg	gat	gcg	tcg	cag	cgt	gtg	acc	caa	cta	ctt	ggc	tct	ctt	act	ata	6211
ser	Asp	Ala	Ser	Gln	Arg	Val	Thr	Gln	Leu	Leu	Gly	ser	Leu	Thr	Ile	
		19	45			1	950				195	5				
acc	agc	cta	ctc	aga	aga	ctc	cac	aat	tgg	ata	act	gag	gac	tgc	ccc	6259
Thr	Ser	Leu	Leu	Arg	Arg	Leu	His	Asn	Trp	lle	Thr	Glu	Asp	Cys	Pro	
	1	960				1965				19	70					
									*							
			tcc													6307
		Cys	Ser	Gly		_	Leu	Arg			Trp	Asp	Trp	Val	Сув	
1	975				198	0			19	85						
														<b>.</b> b. a		6355
			aca Thr													0333
199		пец	1111	199		пур	non		000	1111	ber	2,5	200			
1,,,	•															
aag	ctg	ccc	ggc	ctc	ccc	ttc	atc	tet	tgt	caa	aag	999	tac	aag	ggt	6403
Lys	Leu	Pro	Gly	Leu	Pro	Phe	11e	Ser	Cys	Gln	Lys	Gly	Tyr	Lys	Gly	
			201	0			2	015				2020	)			
gtg	tgg	gcc	ggc	act	ggc	atc	atg	acc	acg	cgc	tgc	cct	tgc	ggc	gcc	6451
Val	тгр	Ala	Gly	Thr	Gly	Ile	Met	Thr	Thr	Arg	Сув	Pro	Сув	Gly	Ala	
		20	25			3	8030				203	5				
aac	atc	tct	ggc	aat	gtc	ege	ctg	ggc	tet	atg	agg	atc	aca	ggg	cct	6499
Asn	Ile	ser	Gly	Asn	Va1	Arg	Leu	Gly	Ser	Met	Arg	Tle	Thr	Gly	Pro	

2040		2045		2050			
aaa acc tgc	atg aac ac	cc tgg cag	ggg acc	ttt cct a	atc aat 1	tgc tac	6547
Lys Thr Cys	Met Asn Th	nr Trp Gln	Gly Thr	Phe Pro	Ile Asn	Cys Tyr	
2055	20	060	20	65			
acg gag ggc	cag tgc gc	cg ccg aaa	ecc ccc	acg aac t	tac aag	acc gcc	6595
Thr Glu Gly	Gln Cys Al	la Pro Lys	Pro Pro	Thr Asn	ryr Lys	Thr Ala	
2070	2075		2080		2085		
atc tgg agg	gtg gcg gc	cc tcg gag	tac gcg	gag gtg a	acg cag	cat ggg	6643
Ile Trp Arg	Val Ala A	la Ser Glu	Tyr Ala	Glu Val	Thr Gln	His Gly	
	2090	2	095	:	2100		
tog tac toc	tat gta ac	ca gga ctg	acc act	gac aat	ctg aaa	att cct	6691
Ser Tyr Ser	Tyr Val T	hr Gly Leu	Thr Thr	Asp Asn	Leu Lys	Ile Pro	
21	05	2110		2115			
tgc caa cta	cct tct cc	ca gag ttt	ttc tcc	tgg gtg	gac ggt	gtg cag	6739
Cys Gln Leu	Pro Ser Pr	ro Glu Phe	Phe Ser	Trp Val	Asp Gly	Val Gln	
2120		2125		2130			
atc cat agg	ttt gca co	cc aca cca	aag ccg	ttt ttc	cgg gat	gag gtc	6787
Ile His Arg	Phe Ala Pr	ro Thr Pro	Lys Pro	Phe Phe	Arg Asp	Glu Val	
2135	2	140	2	145			
tcg ttc tgc	gtt ggg ct	tt aat too	tat gct	gtc ggg l	tcc cag	ctt ccc	6835
Ser Phe Cys	Val Gly Lo	eu Asn Ser	Tyr Ala	Val Gly	Ser Gln	Leu Pro	
2150	2155		2160		2165		
tgt gaa cct	gag ccc ga	ac gca gac	gta ttg	agg tcc	atg cta	aca gat	6883

Cys Glu Pro Glu Pro Asp Ala Asp Val Leu Arg Ser Met Leu Thr Asp

ccg	ccc	cac	atc	açg	gog	gag	act	gcg	gcg	cgg	cgc	ttg	gca	cgg	gga	6931
Pro	Pro	Ris	Ile	Thr	Ala	Glu	Thr	Ala	Ala	Arg	Arg	Leu	Ala	Arg	Gly	
		21	85			:	2190				219	5				
tca	act	cca	tat	gag	gcg	agc	tcc	tca	gtg	age	cag	cta	tca	gca	ccg	6979
Ser	Pro	Pro	ser	Glu	Ala	Ser	Ser	Ser	Val	Ser	Gln	Leu	Ser	Ala	Pro	
	2:	200				2205				22	10					
tog	ctg	cgg	gcc	acc	tgc	acc	acc	cac	agc	aac	acc	tat	gac	gtg	gac	7027
Ser	Leu	Arg	Ala	Thr	Сув	Thr	Thr	His	Ser	Asn	Thr	Tyr	Asp	Val	Asp	
2	215				222	0			22	25						
atg	gtc	gat	gcc	aac	ctg	ctc	atg	gag	ggc	ggt	gtg	gat	cag	aca	gag	7075
Met	Val	Asp	Ala	Asn	Leu	Leu	Met	Glu	Gly	Gly	Va 1	Ala	Gln	Thr	Glu	
223	9			223	5			2	240				224	õ		
cct	gag	tcc	agg	gtg	ccc	gtt	ctg	gac	ttt	ctc	gag	cca	atg	gcc	gag	7123
Pro	Glu	Ser	Arg	Val	Pro	Val	Leu	Asp	Phe	Leu	Glu	Pro	Met	Ala	Glu	
			225	0			2	255				226	9			
gaa	gag	agc	gac	ctt	gag	ccc	tca	ata	cca	teg	gag	tgc	atg	ata	ccc	7171
Glu	Glu	Ser	Asp	Leu	Glu	Pro	Ser	Ile	Pro	Ser	G1 u	Cys	Met	Leu	Pro	
		22	65				2270				227	5				
aqq	age	qqq	ttt	cca	caa	qcc	tta	ccq	gct	tgg	qca	cgg	cct	gac	tac	7219
						Αla										
		80			-	2285				22				•		
aac	cca	cca	ctc	ata	gaa	taa	taa	aαα	agg	cca	σat	tac	caa	ccq	ccc	7267
				5-3		3		- 55								
Asn	Pro	Pro	Leu	Val	Glu	ser	Trp	Arq	Arg	Pro	Asp	Tvr	Gln	Pro	Pro	

acc gtt gct	ggt tgt	get ete	ccc ccc	ccc aag	aag gcc	ccg acg	cct 73	15
Thr Val Ala	Gly Сув	Ala Leu	Pro Pro	Pro Lys	Lys Ala	Pro Thr	Pro	
2310	23	1.5	2	320		2325		
ccc cca agg	aga cgc	cgg aca	gtg ggt	ctg agc	gag agc	acc ata	tca 73	63
Pro Pro Arg	Arg Arg	Arg Thr	Val Gly	Leu Ser	Glu Ser	Thr Ile	ser	
	2330		2335		234	0		
gaa gcc ctc	cag caa	ctg gcc	atc aag	acc ttt	ggc cag	ccc ccc	tog 74	11
Glu Ala Leu	Gln Gln	Leu Ala	Ile Lys	Thr Phe	Gly Gln	Pro Pro	Ser	
23	45		2350		2355			
agc ggt gat	gca ggc	teg tee	acg ggg	gag gga	gaa gaa	gaa tcc	ggc 74	59
Ser Gly Asp	Ala Gly	ser ser	Thr Gly	Ala Gly	Ala Ala	Glu ser	Gly	
2360		236	5	23	70			
ggt ccg acg	tcc cct	ggt gag	acg gad	ccc tca	gag aca	ggt tcc	gcc 75	07
ggt ccg acg Gly Pro Thr							-	07
							-	07
Gly Pro Thr		Gly Glu		Pro Ser			-	07
Gly Pro Thr	Ser Pro	Gly Glu 2380	Pro Ala	Pro Ser 2385	Glu Thr	Gly ser	Ala	
Gly Pro Thr 2375	Ser Pro	Gly Glu 2380 ctc gag	Pro Ala	Pro Ser 2385 det gga	Glu Thr	Gly ser	Ala gag 75	
Gly Pro Thr 2375 tcc tct atg	Ser Pro	Gly Glu 2380 ctc gag Leu Glu	Pro Ala	Pro Ser 2385 det gga	Glu Thr	Gly ser	Ala gag 75	
Gly Pro Thr 2375 tcc tct atg Ser Ser Met	Ser Pro	Gly Glu 2380 ctc gag Leu Glu	Pro Ala	Pro Ser 2385 Cct gga Pro Gly	Glu Thr	Gly ser gac ctg Asp Leu	Ala gag 75	
Gly Pro Thr 2375 tcc tct atg Ser Ser Met	ser Pro	Gly Glu 2380 ctc gag Leu Glu	Pro Ala ggg gag Gly Glu	Pro Ser 2385 det gga Pro Gly	Glu Thr	gac ctg Asp Leu 2405	Ala gag 75 Glu	5 5
Gly Pro Thr 2375 tcc tct atg Ser Ser Met 2390	ser Pro	Gly Glu 2380 ctc gag Leu Glu 95	ggg gag Gly Glu	Pro Ser 2385 Cct gga Pro Gly 2400	gat ccg Asp Pro	gac ctg Asp Leu 2405	gag 75 Glu gct 76	5 5
Gly Pro Thr 2375  tcc tct atg ser Ser Met 2390  tct gat cag	ser Pro	Gly Glu 2380 ctc gag Leu Glu 95	ggg gag Gly Glu	Pro Ser 2385 Cct gga Pro Gly 2400	gat ccg Asp Pro	gac ctg Asp Leu 2405 ggg gta Gly Val	gag 75 Glu gct 76	5 5
Gly Pro Thr 2375  tcc tct atg ser Ser Met 2390  tct gat cag	ser Pro	Gly Glu 2380 ctc gag Leu Glu 95	ggg gag Gly Glu cet ecc Pro Pro	Pro Ser 2385 Cct gga Pro Gly 2400	gat ccg Asp Pro	gac ctg Asp Leu 2405 ggg gta Gly Val	gag 75 Glu gct 76	5 5
Gly Pro Thr 2375  tcc tct atg ser Ser Met 2390  tct gat cag	Ser Pro	Gly Glu 2380  ctc gag Leu Glu 95  ctt caa Leu Gln	ggg gag Gly Glu cet ecc Pro Pro 2415	Pro Ser 2385 det gga Pro Gly 400 dec cag	Glu Thr gat ccg Asp Pro ggg ggg Gly Gly 242	gac ctg Asp Leu 2405 ggg gta Gly Val	gag 75 Glu gct 76	03

. 2295

2425	243	10	2435	
			acc ggg gct cta	
2440	2445	245		
act ccc tgt agc	ece gaa gag ga	a aag ttg cca	atc aac cct ttg	agt 7747
			Ile Asn Pro Leu	
2455	2460	2465		
aac tog otg ttg	cga tac cat aa	ac aag gtg tac	tgt aca aca tca	aag 7795
Asn Ser Leu Leu	Arg Tyr His As	en Lys Val Tyr	Cys Thr Thr Ser	Lys
2470	2475	2480	2485	
			gac agg acg caa	
Ser Ala Ser Gln	Arg Ala Lys Ly		Asp Arg Thr Gln	Val
245	90	2495	2500	
			atc aag cta gcg	
<u>-</u>			Ile Lys Leu Ala	Ala
2505	251	10	2515	
			tan	tta 7939
			gag gcg tgc cag Glu Ala Cys Gln	
-	Ara Arg neu ne	eu ini neu diu		
	2525	25		
2520	2525	25		
			30	
act cca ccc cat	tot goa aga to	cc aag tat gga	ttc ggg gcc aag	gag 7987
act cca ccc cat	tot goa aga to	cc aag tat gga	30	gag 7987
act cca ccc cat	tot goa aga to Ser Ala Arg Se	cc aag tat gga er Lys Tyr Gly	ttc ggg gcc aag	gag 7987
act cca ccc cat Thr Pro Pro His 2535	tot goa aga to Ser Ala Arg Se 2540	cc aag tat gga ar Lys Tyr Gly 2545	ttc ggg gcc aag	gag 7987 Glu

2550	25	55		2560				2569	5		
aag gac ctc	ctg gaa	gac cca	caa ac	a cca	att	ccc	aca	acc	atc	atg	8083
Lys Asp Leu	Leu Glu	Asp Pro	Gln Th	r Pro	Ile	Pro	Thr	Thr	Ile	Met	
	2570		2575	•			2580	)			
gcc aaa aat											8131
Ala Lys Asn	Glu Val	Phe Cys	Val As	p Pro	Ala	ГÀВ	Gly	Gly	Lys	Lys	
2.5	85		2590			259	5				
cca gct cgc											8179
Pro Ala Arg	Leu Ile			p Leu			Arg	Val	Cys	Glu	
2600		260	5		26	10					
aaa atg gcc		_									8227
Lys Met Ala	Leu Tyr	Asp Ile	Thr G1	n Lys	Leu	Pro	GIn	Ala	Val	Met	
		0.000			-0.5						
2615		2620		26	525						
	hab ggg						caa	at a	ana	+++	8275
gga get tee		ttc cag		c cct	gec						8275
gga gct tcc Gly Ala Ser	Tyr Gly	ttc cag		c cct r Pro	gec			Val	Glu		8275
gga get tee		ttc cag		c cct	gec				Glu		8275
gga gct tcc Gly Ala Ser 2630	Tyr Gly	ttc cas Phe Glr	Tyr Se	c cct r Pro 2640	gcc Ala	Gln	Arg	Val 264	Glu 5	Tyr	8275 8323
gga gct tcc Gly Ala Ser 2630	Tyr Gly 26	ttc cag Phe Glr 35	Tyr Se	c cct r Pro 2640 g gac	gcc Ala	Gln atg	Arg ggt	Val 264	Glu 5 tcg	Tyr	
gga gct tcc Gly Ala Ser 2630	Tyr Gly 26	ttc cag Phe Glr 35	Tyr Se	c cct r Pro 2640 g gac s Asp	gcc Ala	Gln atg	Arg ggt	Val 264: ttt Phe	Glu 5 tcg	Tyr	
gga gct tcc Gly Ala Ser 2630	Tyr Gly 26 gca tgg Ala Trp	ttc cag Phe Glr 35	Tyr Se	c cct r Pro 2640 g gac s Asp	gcc Ala	Gln atg	Arg ggt Gly	Val 264: ttt Phe	Glu 5 tcg	Tyr	
gga gct tcc Gly Ala Ser 2630	Tyr Gly 26 gca tgg Ala Trp 2650	ttc cag Phe Glr 35 gcg gaa Ala Glu	aag aa Lys Ly 2655	c cct r Pro 2640 g gac s Asp	gcc Ala ccc Pro	Gln atg Met	ggt Gly 2660	Val 264: ttt Phe	glu tcg ser	Tyr tat Tyr	
gga gct tcc Gly Ala Ser 2630 ctc ttg aaa Leu Leu Lys	Tyr Gly 26  gca tgg Ala Trp 2650  tgc ttc	ttc cag Phe Glr 35 gcg gaa Ala Glu	aag aa Lys Ly 2655	c cct r Pro 2640 g gac s Asp	gcc Ala ccc Pro	Gln atg Met	ggt Gly 266	Val 264: ttt Phe	tcg ser	tat Tyr	8323
gga gct tcc Gly Ala Ser 2630  ctc ttg aaa Leu Leu Lys gat acc cga Asp Thr Arg	Tyr Gly 26  gca tgg Ala Trp 2650  tgc ttc	ttc cag Phe Glr 35 gcg gaa Ala Glu	aag aa Lys Ly 2655	c cct r Pro 2640 g gac s Asp	gcc Ala ccc Pro	Gln atg Met	ggt Gly 2660 gac Asp	Val 264: ttt Phe	tcg ser	tat Tyr	8323
gga gct tcc Gly Ala Ser 2630  ctc ttg aaa Leu Leu Lys gat acc cga Asp Thr Arg	gca tgg Ala Trp 2650  tgc ttc Cys Phe	ttc cag Phe Glr 35 gcg gaa Ala Glu	aag aa Lys Ly 2655	c cct r Pro 2640 g gac s Asp	gcc Ala ccc Pro	Gln atg Met aga Arg	ggt Gly 2660 gac Asp	Val 264: ttt Phe	tcg ser	tat Tyr	8323
gga gct tcc Gly Ala Ser 2630  ctc ttg aaa Leu Leu Lys gat acc cga Asp Thr Arg	Tyr Gly 26 gca tgg Ala Trp 2650 tgc ttc Cys Phe	ttc cas Phe Glr 35 gcg gaa Ala Glu gac tca Asp Ser	aag aa aa bys Ly 2655 acc gt	c cct r Pro 2640 g gac s Asp ; c act	gcc Ala ccc Pro gag	Gln atg Met aga Arg 267	ggt Gly 266 gac Asp	Val 264: ttt Phe 0	tcg ser agg	tat Tyr acc	8323
gga got too Gly Ala Ser 2630  ctc ttg aaa Leu Leu Lys gat acc cga Asp Thr Arg	Tyr Gly 26 gca tgg Ala Trp 2650 tgc ttc Cys Phe 65	ttc cas Phe Glr 35 gcg gaa Ala Glu gac tcs Asp Ser	aag aa aa Lys Ly 2655 acc gt Thr Va 2670	c cct r Pro 2640 g gac s Asp ; ; c act 1 Thr	gcc Ala ccc Pro gag Glu	Gln atg Met aga Arg 267	ggt Gly 2660 gac Asp 5	Val 264: ttt Phe atc Ile	tcg Ser agg Arg	Tyr tat Tyr acc Thr	8323 8371

2680	2685		2690	
			ta gga ggg ccc atg	
2695	2700	270	5	
aac agc aag gg	t caa acc tgc	ggt tac aga c	gt tgc cgc gcc agc	ggg 8515
			rg Cys Arg Ala Ser	Gly
2710	2715	2720	2725	
			es had tak dha asa	gcc 8563
			ca tgc tat gtg aaa hr Cys Tyr Val Lys	
	730	2735	2740	
-				
cta gcg gcc tg	c aag get geg	ggg ata gtt g	eg eec aca atg etg	gta 8611
Leu Ala Ala Cy	s Lys Ala Ala	Gly Ile Val A	la Pro Thr Met Leu	Val
2745	2	750	2755	
tgc ggc gat ga	c cta gta gtc	atc tca gaa a	gc cag ggg act gag	gag 8659
Cys Gly Asp As	p Leu Val Val	Ile Ser Glu S	er Gln Gly Thr Glu	Glu
2760	2765		2770	
			cc atg acc agg tac	
		Phe Thr Glu A	la Met Thr Arg Tyr	ser
2775	2780	278	•	
gec cct cct ga	t gat ccc ccc	aga ccq qaa t	at gac ctg gag cta	ata 8755
-			yr Asp Leu Glu Leu	
2790	2795	2800	2805	
aca too tgt to	c tca aat gtg	tct gtg gcg t	tg ggc ccg cgg ggc	cgc 8803
Thr Ser Cys Se	r Ser Asn Val	Ser Val Ala L	eu Gly Pro Arg Gly	Arg

	2810		28	315			2820	)			
cgc aga tac											8851
Arg Arg Tyr				Pro Thr	Thr			Ala	Arg	Ala	
282	25	2	830			283	5				
gec tgg gaa	aca att	aga cac	tee	cct atc	aat	tca	taa	eta	gga	aac	8899
Ala Trp Glu											
2840		2845			28						
atc atc cag	tat gct	cca acc	ata	tgg gtt	cgc	atg	gtc	cta	atg	aca	8947
Ile Ile Gln	Tyr Ala	Pro Thr	Ile	Trp Val	Arg	Met	Va1	Leu	Met	Thr	
2855		2860		2	865						
cac ttc ttc											8995
His Phe Phe	Ser Ile	Leu Met	Val	Gln Asp	Thr	Leu	Asp	Gln	Asn	Leu	
2870	287	5		2880				288	i		
										abb	0043
aac ttt gag	atg tat	gga tca		tac tec				ttg	gac		9043
	atg tat Met Tyr	gga tca	Val	tac tee			Pro	ttg Leu	gac		9043
aac ttt gag	atg tat	gga tca	Val	tac tec				ttg Leu	gac		9043
aac ttt gag Asn Phe Glu	atg tat Met Tyr 2890	gga tca Gly Ser	Val 28	tac tec Tyr Ser	Val	Asn	Pro 290	ttg Leu	gac Asp	Leu	9043
aac ttt gag	atg tat Met Tyr 2890 att gag	gga tca Gly Ser agg tta	Val 28	tac tec Tyr Ser 895 ggg ctt	Val gac	Asn	Pro 290	ttg Leu )	gac Asp	Leu	
aac ttt gag Asn Phe Glu	atg tat Met Tyr 2890 att gag Ile Glu	gga tca Gly Ser agg tta Arg Leu	Val 28	tac tec Tyr Ser 895 ggg ctt	Val gac	Asn	Pro 290 ttt Phe	ttg Leu )	gac Asp	Leu	
aac ttt gag Asn Phe Glu cca gcc ata Pro Ala Ile	atg tat Met Tyr 2890 att gag Ile Glu	gga tca Gly Ser agg tta Arg Leu	Val 28 cac His	tac tec Tyr Ser 895 ggg ctt	Val gac	Asn gcc Ala	Pro 290 ttt Phe	ttg Leu )	gac Asp	Leu	
aac ttt gag Asn Phe Glu cca gcc ata Pro Ala Ile	atg tat Met Tyr 2890 att gag Ile Glu	gga tca Gly Ser agg tta Arg Leu	Val 28 cac His	tac tec Tyr Ser 395 ggg ett Gly Let	yal gac Asp	gcc Ala 291	Pro 290 ttt Phe 5	teg Leu ) tet Ser	gac Asp atg Met	Leu cac His	
aac ttt gag Asn Phe Glu cca gcc ata Pro Ala Ile	atg tat Met Tyr 2890 att gag Ile Glu 05 cac cac	gga tca Gly Ser agg tta Arg Leu	val  28  cac His  910  acg	tac toc Tyr Ser 395 ggg ctt Gly Let	gac Asp	gcc Ala 291	Pro 290 ttt Phe 5	teg Leu tet Ser	gac Asp atg Met	Leu cac His	9091
aac ttt gag Asn Phe Glu  cca gcc ata Pro Ala Ile 299	atg tat Met Tyr 2890 att gag Ile Glu 05 cac cac	gga tca Gly Ser agg tta Arg Leu	val  28  cac His  2910  acg	tac toc Tyr Ser 395 ggg ctt Gly Let	gac Asp	gcc Ala 291 tca Ser	Pro 290 ttt Phe 5	teg Leu tet Ser	gac Asp atg Met	Leu cac His	9091
aac ttt gag Asn Phe Glu  cca gcc ata Pro Ala Ile 29: aca tac tct Thr Tyr Ser	atg tat Met Tyr 2890 att gag Ile Glu 05 cac cac	gga tca Gly Ser agg tta Arg Leu : gaa ctg	val  28  cac His  2910  acg	tac toc Tyr Ser 395 ggg ctt Gly Let	gac Asp gct Ala	gcc Ala 291 tca Ser	Pro 290 ttt Phe 5	teg Leu tet Ser	gac Asp atg Met	Leu cac His	9091
aac ttt gag Asn Phe Glu  cca gcc ata Pro Ala Ile 29: aca tac tct Thr Tyr Ser	atg tat Met Tyr 2890 att gag Ile Glu 05 cac cac His His	gga tca Gly Ser agg tta Arg Leu gaa ctg Glu Leu 2925	Val 28 cac His 2910 acg Thr	tac tec Tyr Ser 395 ggg ctt Gly Leu cgg gtc Arg Val	gac Asp gct Ala 29	gcc Ala 291 tca Ser 30	Pro 2900 ttt Phe 5 gcc Ala	ttg beu  tot ser  ctc Leu	gac Asp atg Met aga Arg	cac His aaa Lys	9091

2935 2940 2945 agg gcg tcc ctc atc tcc cgt gga ggg aaa gcg gcc gtt tgc ggc cga 9235 Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala Ala Val Cys Gly Arg 2950 2955 2960 2965 tat ctc ttc aat tgg gcg gtg aag acc aag ctc aaa ctc act cca ttg 9283 Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu 2975 2980 2970 ccg gag geg ege eta etg gae tta tee agt tgg tte acc gte gge gec 9331 Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Gly Ala 2995 2985 2990 ggc ggq ggc gac att ttt cac agc gtg tcg cgc cga ccc cgc tca 9379 Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg Ala Arg Pro Arg Ser 3000 3005 3010 tta etc tte gge eta etc eta ett tte gta ggg gta gge etc tte eta 9427 Leu Leu Phe Gly Leu Leu Leu Leu Phe Val Gly Val Gly Leu Phe Leu 3020 3015 3025 ete eee get egg tag ageggeacac actaggtaca etecataget aactgtteet 9482 Leu Pro Ala Arg 3030 tittittt tittittccc 9542 

transportage etgtgaaagg tecqtgagee quatqactgc agagagtqcc gtaactggtc 9662

tototgoaga toatgt 9678

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Thi 35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro

Ile pro Lys Asp Arg Arg Ser Thr Gly Lys Ala Trp Gly Lys Pro Gly 65 70 75 80

Arg Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp

Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro

Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys

Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Val Val Gly Ala Pro Leu 130 135 140

Sor Gly Ala Ala Arg Ala Val Ala His Gly Val Arg Val Leu Glu Asp 145 150 155 160

Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Phe Pro Phe Ser Ile

Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Val Pro Val Ser Ala Ala

		1	В 0				185				19	0			
G1n	Val	Lys	Asn	Thr	Ser	Ser	ser	Tyr	Met	va1	Thr	Asn	Asp	Сув	Ser
	1	95				200				20	5				
Asn	Asp	Ser	Ile	Thr	Trp	Gln	Leu	Glu	Ala	Ala	Val	Leu	His	Va1	Pro
	210				215	i			2	20					
Gly	Cys	Va1	pro	Cys	G1 u	Arg	Va1	Gly	Asn	Thr	ser	Arg	Cys	Trp	Val
225				23	0			2	35				240		
Pro	val	ser	Pro	Asn	Met	A1 a	Val	Arg	Gln	Pro	Gly	Ala	Leu	Thr	Gl n
			24	5			:	250				255			
Gly	Leu	Arg	Thr	His	Ile	Asp	Met	Va1	Val	Met	ser	Ala	Thr	Phe	Сув
		2	60				265				27	0			
ser	Ala	Leu	Tyr	Va1	Gly	Asp	Leu	Сув	G1y	Gly	val	Met	Leu	Ala	Al a
	2	75				280				28	5				
Gln	Va1	Phe	Ile	Val	ser	Pro	Gln	Tyr	His	Trp	Phe	Val	Gln	Glu	Cys
	290				295	;			3	00					
Asn	Cys	Ser	Ile	Tyr	Pro	Gly	Thr	Ile	Thr	Gly	His	Arg	Met	Ala	Trp
305				31	0			1	15				320		
	Met	Met	Met			Ser	Pro			Thr	Met	Ile		Ala	Tyr
	Met	Met	Met 32	Asn		Ser				Thr	Met	11e 339	Leu		Tyr
Asp			32	Asn 5	Trp			Thr 330	Ala			335	Leu		
Asp		Arg	32	Asn 5	Trp			Thr 330	Ala			335 Ser	Leu	Ala	
Asp Val	Met	Arg 3	32 Val 40	Asn 5 Pro	Trp Glu	Val	11e 345	Thr 330 Tle	Ala Asp	īle	Va1 35	335 Ser 0	Leu Gly	Ala	His
Asp Val	Met Gly	Arg 3	32 Val 40	Asn 5 Pro	Trp Glu	Val	11e 345	Thr 330 Tle	Ala Asp	īle	Val 35 Met	335 Ser 0	Leu Gly	Ala Ala	His
Asp Val Trp	Met Gly	Arg 3 Val	32 Val 40 Met	Asn 5 Pro	Trp Glu Gly	Val Leu 360	Ile 345 Ala	Thr 330 Tle Tyr	Ala Asp Phe	Ile Ser	Val 35 Met	335 Ser 0 Gln	Leu Gly Gly	Ala Ala	His Trp
Asp Val Trp	Met Gly	Arg 3 Val	32 Val 40 Met	Asn 5 Pro	Trp Glu Gly	Val Leu 360 Leu	Ile 345 Ala	Thr 330 Tle Tyr	Ala Asp Phe	Ile Ser	Val 35 Met	335 Ser 0 Gln	Leu Gly Gly	Ala Ala Ala	His Trp
Val Trp Ala	Met Gly Lys 370	Arg 3 Val 355 Val	32 Val 40 Met	Asn Fro Pro	Glu Gly Ile 375	Val Leu 360 Leu	Ile 345 Ala Leu	Thr 330 Ile Tyr Leu	Ala Asp Phe Ala 3	Ile Ser 36 Ala 80	Val 35 Met 55 Gly	335 Ser 0 Gln Val	Gly Gly Asp	Ala Ala Ala	His Trp Gly
Val Trp Ala	Met Gly Lys 370	Arg 3 Val 355 Val	32 Val 40 Met	Asn Fro Pro	Glu Gly Ile 375	Val Leu 360 Leu	Ile 345 Ala Leu	Thr 330 Tle Tyr Leu Ala	Ala Asp Phe Ala 3	Ile Ser 36 Ala 80	Val 35 Met 55 Gly	335 Ser 0 Gln Val	Gly Gly Asp	Ala Ala Ala Ala	His Trp Gly
Asp Val Trp Ala Thr 385	Gly Lys 370 Thr	Arg 3 Val 55 Val Thr	32 Val 40 Met Ile Val	Asn Fro Phe Val Gly	Glu Gly Ile 375 Gly	Val Leu 360 Leu Ala	Ile 345 Ala Leu Val	Thr 330 Ile Tyr Leu Ala	Ala Asp Phe Ala Arg	Ser 36 Ala 80 Ser	Val 35 Met 55 Gly Thr	Ser 0 Gln Val	Gly Gly Asp Val	Ala Ala Ala Ala	His Trp Gly Ala
Asp Val Trp Ala Thr 385	Gly Lys 370 Thr	Arg 3 Val 55 Val Thr	32 Val 40 Met Ile Val	Asn Pro Phe Val Gly 39	Glu Gly Ile 375 Gly	Val Leu 360 Leu Ala	Ile 345 Ala Leu Val	Thr 330 Ile Tyr Leu Ala	Ala Asp Phe Ala Arg	Ser 36 Ala 80 Ser	Val 35 Met 55 Gly Thr	Ser 0 Gln Val	Gly Asp Val 400	Ala Ala Ala Ala	His Trp Gly Ala
Asp Val Trp Ala Thr 385	Gly  Lys  370  Thr	3 Val 355 Val Thr	Val 40 Met Ile Val Ser 40	Asn Pro Phe Val Gly 39 His	Glu Gly Ile 375 Gly 0 Gly	Val Leu 360 Leu Ala	Ile 345 Ala Leu Val	Thr 330 Tle Tyr Leu Ala :: Gln 410	Ala Asp Phe Ala 3 Arg 395 Asn	Ile Ser 36 Ala 80 Ser	Val 35 Met 55 Gly Thr	3355 Ser 00 Gln Val Asn Leu	Gly Gly Asp Val 400	Ala Ala Ala Ala	His Trp Gly Ala
Asp Val Trp Ala Thr 385	Gly  Lys  370  Thr	3 Val S55 Val Thr	Val 40 Met Ile Val Ser 40	Asn Pro Phe Val Gly 39 His	Glu Gly Ile 375 Gly 0 Gly	Val Leu 360 Leu Ala	Ile 345 Ala Leu Val	Thr 330 Tle Tyr Leu Ala :: Gln 410	Ala Asp Phe Ala 3 Arg 395 Asn	Ile Ser 36 Ala 80 Ser	Val 35 Met 55 Gly Thr	Ser o Gln Val Asn Leu 415	Gly Gly Asp Val 400	Ala Ala Ala Ile	His Trp Gly Ala

435				440				4.4	5				
Ser Ser Gly	Сув	Pro	Gly	Arg	Leu	ser	Ala	сув	Arg	Asn	Ile	Glu	Ala
450			455	i			4	60					
Phe Arg Ile	Gly	Trp	Gly	Thr	Leu	Gln	Tyr	Glu	Asp	Asn	Val	Thr	Asn
465		47	0			4	75				480		
Pro Glu Asp	Met	Arg	pro	Tyr	Cys	Trp	His	туг	Pro	Pro	Lys	Pro	Cys
	48	5				190				495			
Gly Val Val	Pro	Ala	Arg	Ser	Val	Сув	Gly	Pro	Val	Tyr	Cys	Phe	Thr
5	00				505				51	0			
Pro Ser Pro	<b>V</b> al	Val	Val	Gly	Thr	Thr	Asp	Arg	Arg	Gly	Val	Pro	Thr
515				520				52	25				
Tyr Thr Trp	Gly	Glu	Asn	Glu	Thr	Asp	Val	Phe	Leu	Leu	Asn	Ser	Thr
530			535	j			5	4 0					
Arg Pro Pro	Gln	Gly	Ser	Trp	Phe	Gly	Cys	Thr	Trp	Met	Asn	Ser	Thr
545		55	0				555				560		
Gly Phe Thr	Lvs	Thr	Cvs	Glv	Ala	Pro	pro	Сув	Arg	Thr	Arg	Ala	Asp
01, 1110 1111	2,0	****	0,0						-				
01, 110 111	56		0,0			570				575			
Phe Asn Ala	56	5				570				575	i		
Phe Asn Ala	56	5				570				575 Cys	i		
Phe Asn Ala	56 Ser 80	5 Thr	Asp	Leu	Leu 585	570 Cys	Pro	Thr	Asp 59	575 Cys 0	Phe	Arg	Lys
Phe Asn Ala	56 Ser 80	5 Thr	Asp	Leu	Leu 585	570 Cys	Pro	Thr	Asp 59 Gly	575 Cys 0	Phe	Arg	Lys
Phe Asn Ala 5 His Pro Asp	56 Ser 80 Ala	5 Thr Thr	Asp Tyr	Leu Ile 600	Leu 585 Lys	570 Сув Сув	Pro	Thr Ser	Asp 59 Gly 5	575 Cys 0 Pro	Phe Trp	Arg	Lys Thr
Phe Asn Ala 5 His Pro Asp 595	56 Ser 80 Ala	5 Thr Thr	Asp Tyr	Leu Ile 600 Tyr	Leu 585 Lys	570 Сув Сув	Pro Gly Arg	Thr Ser	Asp 59 Gly 5	575 Cys 0 Pro	Phe Trp	Arg	Lys Thr
Phe Asn Ala 5 His Pro Asp 595 Pro Lys Cys	56 Ser 80 Ala Leu	5 Thr Thr Val	Asp Tyr His	Leu Ile 600 Tyr	Leu 585 Lys Pro	Cys Cys Cys	Pro Gly Arg	Thr Ser 60 Leu 20	Asp 59 Gly 5 Trp	575 Cys 0 Pro	Phe Trp Tyr	Arg Leu Pro	Lys Thr Cys
Phe Asn Ala 5 His Pro Asp 595 Pro Lys Cys 610	56 Ser 80 Ala Leu	5 Thr Thr Val	Asp Tyr His 619	Leu Ile 600 Tyr	Leu 585 Lys Pro	Cys Cys Tyr	Pro Gly Arg	Thr Ser 60 Leu 20	Asp 59 Gly 5 Trp	575 Cys 0 Pro	Phe Trp Tyr	Arg Leu Pro	Lys Thr Cys
Phe Asn Ala 5 His Pro Asp 595 Pro Lys Cys 610 Thr Val Asn	56 Ser 80 Ala Leu	5 Thr Thr Val Thr	Asp Tyr His 619 Tle	Ile 600 Tyr	Leu 585 Lys Pro	Cys Cys Tyr	Pro Gly Arg 6 Arg	Thr Ser 60 Leu 20 Met	Asp 59 Gly 5 Trp Tyr	575 Cys 0 Pro His	Phe Trp Tyr Gly	Arg Leu Pro	Lys Thr Cys
Phe Asn Ala 5 His Pro Asp 595 Pro Lys Cys 610 Thr Val Asn 625	56 Ser 80 Ala Leu	Thr Thr Val Thr 63	Asp Tyr His 619 Tle	Ile 600 Tyr	Leu 585 Lys Pro Lys	Cys Cys Tyr	Pro Gly Arg 6 Arg	Thr Ser 60 Leu 20 Met	Asp 59 Gly 5 Trp Tyr	575 Cys 0 Pro His	Phe Trp Tyr Gly 640	Arg Leu Pro	Lys Thr Cys
Phe Asn Ala 5 His Pro Asp 595 Pro Lys Cys 610 Thr Val Asn 625	56 Ser 80 Ala Leu Phe Leu	5 Thr Thr Val Thr 63 Thr	Asp Tyr His 61: Tle 0 Ala	Ile 600 Tyr Phe	Leu 585 Lys Pro Lys	Cys Cys Tyr Ile Asn	Pro Gly Arg 6 Arg 535	Thr Ser 60 Leu 20 Met	Asp 59 Gly D5 Trp Tyr	Cys O Pro His Val	Trp Tyr Gly 640 Asp	Arg  Leu  Pro  Gly	Lys Thr Cys Val
Phe Asn Ala 5 His Pro Asp 595 Pro Lys Cys 610 Thr Val Asn 625 Glu His Arg	56 Ser 80 Ala Leu Phe Leu	5 Thr Thr Val Thr 63 Thr	Asp Tyr His 61: Tle 0	Ile 600 Tyr Phe	Leu 585 Lys Pro Lys	Cys Cys Tyr Ile Asn	Pro Gly Arg 6 Arg 535	Thr Ser 60 Leu 20 Met	Asp 59 Gly D5 Trp Tyr	Cys Cys O Pro His Val Gly 65:	Trp Tyr Gly 640 Asp	Arg  Leu  Pro  Gly	Lys Thr Cys Val
Phe Asn Ala 5 His Pro Asp 595 Pro Lys Cys 610 Thr Val Asn 625 Glu His Arg	566 Ser 80 Ala Leu Phe Leu 64 Asp	5 Thr Val Thr 63 Thr 5	Asp Tyr His 619 Ile 0 Ala Asp	Ile 600 Tyr Phe Ala	Leu 585 Lys Pro Lys Cys Ser 665	Cys Cys Tyr Ile Asn 650 Gln	Pro Gly Arg 6 Arg Fig. 35 Phe	Thr Ser 66 Leu 20 Met Thr	Asp 59 Gly D5 Trp Tyr Arg	Cys Cys O Pro His Val Gly 659 Leu	Phe Trp Tyr Gly 640 Asp	Arg Leu Pro Gly Arg	Lys Thr Cys Val Cys
Phe Asn Ala  5 His Pro Asp 595 Pro Lys Cys 610 Thr Val Asn 625 Glu His Arg Asp Leu Glu 6	Ser 80 Ala Leu Phe 64 Asp 60 Trp	5 Thr Thr Val Thr 63 Thr 5 Arg	Asp Tyr His 61: Tle 0 Ala Asp	Leu Tle 600 Tyr Phe Ala Arg	Leu 585 Lys Pro Lys Cys Ser 665	Cys Cys Tyr Ile Asn 650 Gln Cys	Pro Gly Arg 6 Arg 535 Phe Leu	Thr Ser 60 Leu 20 Met Thr Ser	Asp 59 Gly 05 Trp Tyr Arg Pro 67 Ser	Cys Cys O Pro His Val Gly 655 Leu O Asp	Trp Tyr Gly 640 Asp Leu	Arg Leu Pro Gly Arg	Lys Thr Cys Val Cys Ser

	690				695				7	0 0					
Tyr	Met	Tyr	Gly	Leu	Ser	Pro	Ala	Ile	Thr	Lys	Tyr	Val	Val	Arg	Trp
705				71	0			7	15				720		
Glu	Trp	Val	Val	Leu	Leu	Phe	Leu	Leu	Leu	Ala	Авр	Ala	Arg	va1	Cys
			72	5				730				735			
Ala	Сув	Leu	Trp	Met	Leu	Ile	Leu	Leu	Gly	Gln	Ala	Glu	Ala	Ala	Leu
		7	40				745				75	0			
Glu	Lys	Leu	Va1	Val	Leu	нів	Ala	Ala	Ser	Ala	Ala	Asn	Cys	His	Gly
	7	55				760				76	5				
Leu	Leu	Tyr	Phe	Ala	11e	Phe	Phe	Val	Ala	Ala	Trp	His	Ile	Arg	Gly
	770				775	;			7	80					
Arg	Val	Val	Pro	Leu	Thr	Thr	Tyr	Сув	Leu	Thr	Gly	Leu	Trp	Pro	Phe
785				79	0			7	195				800		
Сув	Leu	Leu	Leu	Met	Ala	Leu	Pro	Arg	Gln	Ala	Tyr	Ala	Tyr	Asp	Ala
			8 0	15			;	810				815			
Pro	Va 1	His	Gly	Gln	Ile	Gly	<b>Val</b>	Gly	Leu	Leu	Ile	Leu	Ile	Thr	Leu
		8	20				825				83	0			
Phe	Thr	Leu	Thr	Pro	Gly	Tyr	Lys	Thr	Leu	Leu	Gly	Gln	Сув	Leu	Trp
	8	3 5				840				84	5				
Trp	Leu	Сув	Tyr	Leu	Leu	Thr	Leu	Gly	Glu	Ala	Met	Ile	Gln	Glu	Trp
	850				855	i			8	60					
Val	Pro	Pro	Met	Gln	Va 1	Arg	Gly	Gly	Arg	Asp	Gly	Ile	Ala	Trp	Ala
865				87	0			8	375				880		
Val	Thr	lle	Phe	Сув	Pro	Gly	Val	Val	Phe	Asp	Ile	Thr	Гла	тгр	Leu
			88	5				890				895			
Leu	Ala	Leu	Leu	Gly	Pro	Ala	Tyr	Leu	Leu	Arg	Ala	Ala	Leu	Thr	His
		9	00				905				91	0			
val	Pro	Tyr	Phe	Va1	Arg	Ala	ніѕ	Ala	Leu	Ile	Arg	Val	Суя	Ala	Leu
	9	15				920				92	5				
Val	Lys	G1n	Leu	Ala	Gly	Gly	Arg	Tyr	Val	Gln	va1	Ala	Leu	Leu	Ala
	930				935	i			9	40					
					_					Asp	***	×	m1		

945			95	0				955				960		
Ser Asp	Trp	Ala	Ala	ser	Gly	Leu	Arg	Asp	Leu	Ala	Val	Ala	Val	Glu
		96	5				970				975			
Pro Ile	11e	Phe	Ser	Pro	мet	Glu	Lys	ьуз	val	Ile	Val	Trp	Gly	Ala
	9	в 0				985				99	0			
Glu Thr	Ala	Ala	Cys	Gly	Asp	Ile	Leu	His	Gly	Leu	Pro	Va 1	Ser	Ala
9	95				1000				100	15				
Arg Leu	Gly	Gln	Glu	Ile	Leu	Leu	Gly	Pro	Ala	qaA	Gly	Tyr	Thr	Ser
1010				101	5			10	20					
Lys Gly	Trp	Lys	Leu	Leu	Ala	Pro	Ile	Thr	Ala	Tyr	Ala	Gln	Gln	Thr
1025			103	0			1	035				104	D	
Arg Gly	Leu	Leu	Gly	Ala	Ile	Val	Val	Ser	Met	Thr	Gly	Arg	Asp	Arg
		104	5			1	050				105	5		
Thr Glu	Gln	Ala	Gly	Glu	Val	Gln	Ile	Leu	Ser	Thr	Val	Ser	Gln	Ser
	10	60			:	1065				107	0			
Phe Leu	Glv	Thr	Thr	Ile	Ser	Gly	Val	Leu	Trp	Thr	Val	Tyr	His	Gly
	2													
	75				1080				10					
	75				1080			Arg		85				
10	75				1080 Ala					85				
10 Ala Gly	75 Asn	Lys	Thr	Leu 109	1080 Ala 5	Gly	Leu	11	G1y 100	85 Pro	Val	Thr	Gln	Met
10 Ala Gly 1090	75 Asn	Lys	Thr	Leu 109 Gly	1080 Ala 5	Gly	Leu Val	11	G1y 100	85 Pro	Val	Thr	Gln Pro	Met
Ala Gly 1090 Tyr Ser	75 Asn Ser	Lys Ala	Thr Glu 111	Leu 109 Gly	1080 Ala 5 Asp	Gly Leu	Leu Val	11 Gly 115	Gly LOO Trp	Pro	Val Ser	Thr Pro	Gln Pro	Меt Gly
Ala Gly 1090 Tyr Ser 1105	75 Asn Ser	Lys Ala	Thr Glu 111	Leu 109 Gly	1080 Ala 5 Asp	Gly Leu Lys	Leu Val	11 Gly 115	Gly LOO Trp	Pro	Val Ser	Thr Pro 112 Leu	Gln Pro	Меt Gly
Ala Gly 1090 Tyr Ser 1105	75 Asn Ser	Lys Ala Leu 112	Glu 111 Glu 61u	Leu 109 Gly 0 Pro	Ala 5 Asp Cys	Gly Leu Lys 1	Leu Val 1 Cys 130	Gly 115 Gly	Gly LOO Trp	Pro Pro Val	Val Ser Asp	Thr Pro 112 Leu	Gln Pro O Tyr	Met Gly Leu
Ala Gly 1090 Tyr Ser 1105 Thr Lys	75 Asn Ser	Lys Ala Leu 112 Asn	Glu 111 Glu 61u	Leu 109 Gly 0 Pro	Ala 5 Asp Cys	Gly Leu Lys 1	Leu Val 1 Cys 130	Gly 115 Gly	Gly LOO Trp	Pro Pro Val	Val Ser Asp 113:	Thr Pro 112 Leu	Gln Pro O Tyr	Met Gly Leu
Ala Gly 1090 Tyr Ser 1105 Thr Lys	75 Asn Ser Ser Arg	Lys Ala Leu 112 Asn 40	Glu 111 Glu 55 Ala	Leu 109 Gly 0 Pro	Ala  Asp  Cys  Val	Gly Leu Lys 1 Ile	Val 1 Cys 130 Pro	Gly 115 Gly Ala	Gly LOO Trp Ala	Pro Pro Val Arg	Val Ser Asp 113: Arg	Thr Pro 112 Leu 5	Gln Pro Tyr Asp	Met Gly Leu Lys
Ala Gly 1090 Tyr Ser 1105 Thr Lys Val Thr	75 Asn Ser Ser Arg	Lys Ala Leu 112 Asn 40	Glu 111 Glu 55 Ala	Leu 109 Gly 0 Pro	Ala  Asp  Cys  Val	Gly Leu Lys 1 Ile 1145	Val 1 Cys 130 Pro	Gly 115 Gly Ala	Gly LOO Trp Ala	Pro Pro Val Arg 115	Val Ser Asp 113: Arg	Thr Pro 112 Leu 5	Gln Pro Tyr Asp	Met Gly Leu Lys
Ala Gly 1090 Tyr Ser 1105 Thr Lys Val Thr	Ser Ser Arg 11 Ala	Lys Ala Leu 112 Asn 40 Leu	Thr Glu 111 Glu 25 Ala	Leu 109 Gly 0 Pro	Ala  Asp  Cys  Val  Pro  1160	Gly Leu Lys 1 Ile Arg	Val 1 Cys 130 Pro	Gly 115 Gly Ala Tle	Gly 100 Trp Ala Arg Ser	Pro Val Arg 115 Thr	Val Ser Asp 1133 Arg 60	Thr Pro 112 Leu 5	Gln Pro Tyr Asp	Met Gly Leu Lys Ser
Ala Gly 1090 Tyr Ser 1105 Thr Lys Val Thr Arg Gly	Ser Ser Arg 11 Ala	Lys Ala Leu 112 Asn 40 Leu	Thr Glu 111 Glu 25 Ala	Leu 109 Gly 0 Pro	Ala  Asp  Cys  Val  Pro  1160  Cys	Gly Leu Lys 1 Ile Arg	Val 1 Cys 130 Pro	Gly 115 Gly Ala Ile	Gly 100 Trp Ala Arg Ser	Pro Val Arg 115 Thr	Val Ser Asp 1133 Arg 60	Thr Pro 112 Leu 5	Gln Pro Tyr Asp	Met Gly Leu Lys Ser
Ala Gly 1090 Tyr Ser 1105 Thr Lys Val Thr Arg Gly 11 Ser Gly	Ser Ser Arg 11 Ala 55	Lys Ala Leu 112 Asn 40 Leu	Glu 111) Glu 25 Ala Leu Val	Leu 109 Gly 0 Pro Asp Ser Leu 117	Ala  Asp  Cys  Val  Pro  1160  Cys	Lys 1 1le 1145 Arg	Val 1 Cys 130 Pro	Gly 115 Gly Ala Ile Gly 1:	Gly 100 Trp Ala Arg Ser 11 His	Pro Val Arg 115 Thr 65 Val	Val Ser Asp 1133 Arg 60 Leu Val	Thr Pro 112 Leu 5 Gly Lys	Pro Tyr Asp	Met Gly Leu Lys Ser
Ala Gly 1090 Tyr Ser 1105 Thr Lys Val Thr Arg Gly 1170	Ser Ser Arg 11 Ala 55	Lys Ala Leu 112 Asn 40 Leu	Glu 111) Glu 25 Ala Leu Val	Leu 109 Gly 0 Pro Asp Ser Leu 117 Ser	Ala  Asp  Cys  Val  Pro  1160  Cys	Lys 1 1le 1145 Arg	Val Cys 130 Pro Pro Arg	Gly 115 Gly Ala Ile Gly 1:	Gly 100 Trp Ala Arg Ser 11 His	Pro Val Arg 115 Thr 65 Val	Val Ser Asp 1133 Arg 60 Leu Val	Thr Pro 112 Leu 5 Gly Lys	Gln Pro Tyr Asp Gly Leu	Met Gly Leu Lys Ser

			120	5			1	210				121	5		
Asn	Ser	Thr	Pro	Pro	Ala	Val	Pro	Gln	Thr	Tyr	Gln	Val	Gly	Tyr	Leu
		12	20			1	.225				123	0			
His	Ala	Pro	Thr	Gly	Ser	Gly	Lys	Ser	Thr	Lys	Va1	Pro	Val	Ala	Tyr
	12	35				1240				12	15				
Ala	Ala	Gln	Gly	Tyr	Lys	Val	Leu	Val	Leu	Asn	Pro	ser	Val	Ala	Λla
1	250				125	5			12	60					
Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Leu	Ser	Lys	Ala	His	Gly	Ile	Asn	Pro
1265	5			127	0			1	275				1280	)	
Asn	Ile	Arg	Thr	Gly	Val	Arg	Thr	Val	Met	Thr	Gly	Glu	Ala	Ile	Thr
			128	35			1	290				129	5		
Tyr	ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala	Asp	Gly	Gly	Cys	Ala	ser	Gly
		13	00			:	305				131	0			
Ala	Tyr	Авр	Ile	Ile	Ile	Сув	Asp	Glu	Cys	His	Ala	Val	Asp	Ala	Thr
	13	15				1320				13	25				
ser	Ile	Leu	Gly	Tle	Gly	Thr	Val	Leu	Asp	Gln	Ala	Glu	Thr	Ala	Gly
1	330				133	5			13	40					
Val	Arg	Leu	Thr	Val	Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	val	Thr
1345	5			135	50			1	355				136	0	
Thr							a1	17-1			a1	3-0			G311
	Pro	His	Pro	Asp	iie	GIU	GIU	val	GIA	Leu	GLY	MIG	GIU	Gly	GIU
	Pro	His	Pro 136		116	GIU		370	GIÀ	Leu	GLY	137		Gly	gru
Ile			136				1	370	Ī		Ī	137	5		
Ile			136 Tyr	55		Ala	1	370	Ī		Ī	137	5		
	Pro	Phe	136 Tyr 80	55	Arg	Ala	1 Ile 385	370 Pro	Leu	Ser	Cys 139	137 Ile	5 Lys	Gly	Gly
	Pro His	Phe	136 Tyr 80	Gly	Arg	Ala	Ile 385 Ser	370 Pro	Leu	Ser	Сув 139 Сув	137 Ile	5 Lys	Gly	Gly
Arg	Pro His	Phe 13 Leu	136 Tyr 80 Ile	Gly	Arg Cys	Ala : His	Ile 385 Ser	370 Pro Lys	Leu Lys	Ser Lys	Сув 139 Сув 05	Ile 0 Asp	Lys Glu	Gly Leu	Gly Ala
Arg Ala	Pro His	Phe 13 Leu	136 Tyr 80 Ile	Gly Phe	Arg Cys	Ala His 1400 Gly	Ile 385 Ser	370 Pro Lys	Leu Lys Ala	Ser Lys	Сув 139 Сув 05	Ile 0 Asp	Lys Glu	Gly Leu	Gly Ala
Arg Ala	Pro His 13 Ala 410	Phe 13 Leu 395 Leu	Tyr 80 Ile Arg	Gly Phe	Arg Cys Met	Ala His 1400 Gly	Ile 385 Ser Leu	370 Pro Lys Asn	Leu Lys Ala	Ser Lys 14 Val	Cys 139 Cys 05 Ala	137: Ile 0 Asp	<b>Lys</b> Glu Tyr	Gly Leu Arg	Gly Ala Gly
Arg Ala	Pro His 13 Ala 410	Phe 13 Leu 395 Leu	Tyr 80 Ile Arg	Gly Phe	Arg Cys Met 141	Ala His 1400 Gly 5	Ile 385 Ser Leu	370 Pro Lys Asn Gln	Leu Lys Ala	Ser Lys 14 Val	Cys 139 Cys 05 Ala	137: Ile 0 Asp	<b>Lys</b> Glu Tyr	Gly Leu Arg Val	Gly Ala Gly
Arg Ala 1 Leu 1425	Pro His 13 Ala 410 Asp	Phe 13 Leu 395 Leu Val	Tyr 80 Ile Arg	Gly Phe Gly	Arg Cys Met 141 Ile	Ala His 1400 Gly Fro	Ile 385 Ser Leu	370 Pro Lys Asn Gln	Leu Lys Ala 14 Gly 435	Lys 14 Val 20 Asp	Cys 139 Cys 05 Ala Val	1379 Ile O Asp Tyr	Lys Glu Tyr Val	Gly Leu Arg Val	Gly Ala Gly Ala
Arg Ala 1 Leu 1425	Pro His 13 Ala 410 Asp	Phe 13 Leu 395 Leu Val	Tyr 80 Ile Arg	Gly Phe Gly Ile 143	Arg Cys Met 141 Ile	Ala His 1400 Gly Fro	1 Ile 385 Ser Leu Ala	370 Pro Lys Asn Gln	Leu Lys Ala 14 Gly 435	Lys 14 Val 20 Asp	Cys 139 Cys 05 Ala Val	1379 Ile O Asp Tyr	Lys Glu Tyr Val 1440 Ser	Gly Leu Arg Val	Gly Ala Gly Ala

		14	60			1	465				147	0			
Thr	Phe	Thr	Ile	Thr	Thr	Gln	Thr	Val	Pro	Gln	Авр	Ala	Val	Ser	Arg
	14	75				1480				148	35				
Ser	Gln	Arg	Arg	Gly	Arg	Thr	Gly	Arg	Gly	Arg	Gln	Gly	Thr	Tyr	Arg
1	490				149	ö			15	00					
Tyr	Val	Ser	Thr	Gly	Glu	Arg	Ala	ser	Gly	Met	Phe	Asp	Ser	Val	Val
150	5			151	0			1	515				152	0	
Leu	Суз	Glu	Сув	Tyr	qeA	Ala	Gly	Ala	Ala	Trp	туr	Asp	Leu	Thr	Pro
			152	5			1	530				1535	5		
Ala	Glu	Thr	Thr	va 1	Arg	Leu	Arg	Ala	Tyr	Phe	Asn	Thr	Pro	Gly	Leu
		15	40			1	545				155	0			
Pro	Va 1	Сув	Gln	Asp	His	Leu	Glu	Phe	Trp	Glu	Ala	Val	Phe	Thr	Gly
	1	555				1560				15	65				
Leu	Thr	His	Ile	Asp	Ala	His	Phe	Leu	Ser	Gln	Thr	Lys	Gln	Ala	Gly
1	570				157	5			15	80					
Glu	Asn	Phe	Ala	Tyr	Leu	Val	Ala	Tyr	Gln	Ala	Thr	Val	Cys	Ala	Arg
158	5			159	0			1	595				160	0	
		Ala	Pro	159 Pro	-	ser	тгр			Met	Trp	Lys			Ala
		Ala	Pro	Pro	-	Ser				Met	Trp	Lys 161	Сув		Ala
Ala	Lys		160	Pro	Pro		1	Asp 610	Ala			161	Сув 5	Leu	
Ala	Lys	Lys	160	Pro	Pro	Ala	1	Asp 610	Ala			161 Leu	Сув 5	Leu	
Ala	Lys	Lys 16	160 Pro 20	Pro	Pro	Ala	1 Gly 1625	Asp 610 Pro	Ala Thr	Pro	Leu 163	161 Leu 10	Cys 5 Tyr	Leu	Leu
Ala	Lys Leu Pro	Lys 16	160 Pro 20	Pro 5 Thr	Pro	Ala	Gly 1625 Thr	Asp 610 Pro	Ala Thr	Pro	Leu 163 Pro	161 Leu 10	Cys 5 Tyr	Leu	Leu
Ala Arg	Lys Leu Pro	Lys 16 Ile 635	Pro 20 Thr	Pro 5 Thr	Pro Leu Glu	Ala : Val 1640	Gly 1625 Thr	Asp 610 Pro Leu	Ala Thr	Pro His	Leu 163 Pro 45	161 Leu 30 Gly	Cys 5 Tyr Thr	Leu Arg Lys	Leu Tyr
Ala Arg Gly	Lys Leu Pro	Lys 16 Ile 635	Pro 20 Thr	Pro 5 Thr	Pro Leu Glu	Ala Val 1640 Ala	Gly 1625 Thr	Asp 610 Pro Leu	Ala Thr Thr	Pro His	Leu 163 Pro 45	161 Leu 30 Gly	Cys 5 Tyr Thr	Leu Arg Lys	Leu Tyr
Ala Arg Gly Ile	Lys Leu Pro 1 Ala	Lys 16 Ile 635 Thr	Pro 20 Thr	Pro 5 Thr	Pro Leu Glu Gln 165	Ala : Val 1640 Ala	Gly 1625 Thr	Asp 610 Pro Leu Leu	Thr Thr Glu	Pro His 16 Val	Leu 163 Pro 45 Met	161 Leu 30 Gly Thr	Cys 5 Tyr Thr	Leu Arg Lys	Leu Tyr Trp
Ala Arg Gly Ile	Lys Leu Pro 1 Ala 650	Lys 16 Ile 635 Thr	Pro 20 Thr	Pro 5 Thr Asn	Pro Leu Glu Gln 165 Val	Ala : Val 1640 Ala	Gly 1625 Thr	Asp 610 Pro Leu Leu	Thr Thr Glu	Pro His 16 Val	Leu 163 Pro 45 Met	161 Leu 30 Gly Thr	Cys 5 Tyr Thr	Leu Arg Lys Thr	Leu Tyr Trp
Ala Arg Gly Ile 1 Val	Lys Leu Pro 1: Ala .650 Leu 5	Lys 16 Ile 635 Thr	Pro 20 Thr Cys	Pro Thr Asn Met	Pro Leu Glu Gln 165 Val	Ala Val 1640 Ala 5 Leu	Gly L625 Thr Asp	Asp 610 Pro Leu Leu Ala	Thr Thr Glu 16 Val 675	Pro His 16 Val 560	Leu 163 Pro 45 Met	161 Leu 30 Gly Thr	Tyr Thr Ser Cys	Leu Arg Lys Thr Leu	Leu Tyr Trp
Ala Arg Gly Ile 1 Val	Lys Leu Pro 1: Ala .650 Leu 5	Lys 16 Ile 635 Thr	Pro 20 Thr Cys	Pro D5 Thr Asn Met Gly 167 Ser	Pro Leu Glu Gln 165 Val	Ala Val 1640 Ala 5 Leu	Gly L625 Thr Asp Ala	Asp 610 Pro Leu Leu Ala	Thr Thr Glu 16 Val 675	Pro His 16 Val 560	Leu 163 Pro 45 Met	161 Leu 30 Gly Thr	Cys 5 Tyr Thr Ser Cys 168	Leu Arg Lys Thr Leu	Leu Tyr Trp
Ala Arg Gly Ile 1 Val 1666	Lys Leu Pro 10 Ala 650 Leu 5	Lys 16 11e 635 Thr Ala	Pro 20 Thr Cys Gly Val	Pro	Pro Leu Glu Gln 165 Val	Ala : Val 1640 Ala 5 Leu	1 Gly 1625 Thr Asp Ala Gly	Asp 610 Pro Leu Leu Ala 1 Arg	Thr Thr Glu 16 Val 675 Leu	Pro His 16 Val 560 Ala	Leu 163 Pro 45 Met Ala	161 Leu 30 Gly Thr Tyr Asn 169	Cys 5 Tyr Thr Ser Cys 168 Gln	Lys Thr Leu 0	Leu Tyr Trp
Ala Arg Gly Ile 1 Val 1666	Lys Leu Pro 10 Ala 650 Leu 5	Lys 16 11e 635 Thr Ala Cys	Pro 20 Thr Cys Gly Val	Pro	Pro Leu Glu Gln 165 Val	Ala  Val  1640  Ala  5  Leu  Ile	1 Gly 1625 Thr Asp Ala Gly	Asp 610 Pro Leu Leu Ala 1 Arg	Thr Thr Glu 16 Val 675 Leu	Pro His 16 Val 560 Ala	Leu 163 Pro 45 Met Ala	161 Leu 30 Gly Thr Tyr Asn 169	Cys 5 Tyr Thr Ser Cys 168 Gln	Lys Thr Leu 0	Leu Tyr Trp Ala Val

1715			1720			177	25				
Ala Glu Met	Leu	Lys Ser	Lys I	le Gln	Gly	Leu	Leu	Gln	Gln	Ala	Ser
1730		173	5		17	40					
Lys Gln Ala	G1n	Asp Ile	Gln P	ro Ala	Met	Gln	Ala	Ser	Trp	Pro	Lys
1745		1750		1	755				176	)	
Val Glu Glr	Phe	Trp Ala	Arg H	is Met	Trp	Asn	Phe	пе	Ser	Gly	Ile
	176	5		1770				177	5		
Gln Tyr Lev	Ala	Gly Leu	Ser T	thr Leu	Pro	Gly	Asn	Pro	Ala	Va1	Ala
1	780		17	85			179	0			
Ser Met Met	Ala	Phe Ser	Ala A	Ala Leu	Thr	Ser	Pro	Leu	ser	Thr	Ser
1795			1800			180	5				
Thr Thr Ile	Leu	Leu Asn	Ile M	et Gly	Gly	Trp	Leu	Ala	ser	Gln	Ile
1810		181	5		18	20					
Ala Pro Pro	Ala	Gly Ala	Thr G	ly Phe	Val	Val	Ser	Gly	Leu	Val	Gly
1825		1830		1	835				184	)	
Ala Ala Val	Gly	Ser Ile	Gly L	eu Gly	Lys	val	Leu	Val	Asp	Ile	Leu
	184			1850				185	5		
Ala Gly Tyr	184	5	Ile S	1850			Val			Lys	Ile
	184	5		1850			Val 187	Ala		Lys	Ile
	184 Gly 860	5 Ala Gly	18	1850 Ser Gly	λla	Leu	187	Ala 0	Phe		
1	184 Gly 860	5 Ala Gly Lys Pro	18	1850 Ser Gly	λla	Leu	187 Ile	Ala 0	Phe		
1 Met Ser Gly	184 Gly 860 Glu	5 Ala Gly Lys Pro	18 Ser M 1880	1850 Ger Gly 865 Met Glu	Ala Asp	Leu Val 18	187 Ile 35	Ala 0 Asn	Phe	Leu	Pro
1 Met Ser Gly 1875	184 Gly 860 Glu	5 Ala Gly Lys Pro	18 Ser M 1880 Ala L	1850 Ger Gly 865 Met Glu	Ala Asp	Leu Val 18: Gly	187 Ile 35	Ala 0 Asn	Phe	Leu	Pro
1 Met Ser Gly 1875 Gly Ile Let	184 Gly 860 Glu	5 Ala Gly Lys Pro Pro Gly 189	18 Ser M 1880 Ala L	1850 Ger Gly 865 Met Glu Leu Val	Ala Asp Val	Leu Val 18: Gly	187 Ile 85 Val	Ala 0 Asn Ile	Phe Leu Cys	Leu Ala	Pro Ala
1 Met Ser Gly 1875 Gly Ile Lev 1890	184 Gly 860 Glu	5 Ala Gly Lys Pro Pro Gly 189	18 Ser M 1880 Ala L	1850 Ser Gly 865 Met Glu Leu Val	Ala Asp Val	Leu Val 18: Gly	187 Ile 85 Val	Ala 0 Asn Ile	Phe Leu Cys	Leu Ala Trp	Pro Ala
1 Met Ser Gly 1875 Gly Ile Leu 1890 Ile Leu Arg	184 Gly 860 Glu ser	5 Ala Gly Lys Pro Pro Gly 189 His Val	18 Ser M 1880 Ala L 5	1850 Ger Gly 365 Met Glu Geu Val	Ala Asp Val 19 Glu 915	Leu Val 18: Gly 00 Gly	187 Ile 35 Val Ala	Ala  O  Asn  Ile  Val	Phe Leu Cys Gln 192	Leu Ala Trp	Pro Ala Met
Met Ser Gly 1875 Gly Ile Leu 1890 Ile Leu Arg	184 Gly 860 Glu ser	5 Ala Gly Lys Pro Pro Gly 189 His Val 1910 Ala Phe	18 Ser M 1880 Ala L 5	1850 Ger Gly 365 Met Glu Geu Val	Ala Asp Val 19 Glu 915	Leu Val 18: Gly 00 Gly	187 Ile 35 Val Ala	Ala  O  Asn  Ile  Val	Phe Leu Cys Gln 192	Leu Ala Trp	Pro Ala Met
Met Ser Gly 1875 Gly Ile Leu 1890 Ile Leu Arg	184: Gly 360 Glu Ser Arg	5 Ala Gly Lys Pro Pro Gly 189: His Val 1910 Ala Phe	18 Ser M 1880 Ala L 5 Gly P Ala S	1850 Ser Gly 665 Met Glu Geu Val Pro Gly 1 Ser Arg	Ala Asp Val 19 Glu 915	Val 188 Gly 00 Gly Asn	187 Ile 85 Val Ala	Ala  O  Asn  Ile  Val  Val  193	Phe Leu Cys Gln 1920 Ala	Leu Ala Trp Pro	Pro Ala Met
1875 Gly Ile Let 1890 Ile Leu Arg 1905 Asn Arg Let	184: Gly 360 Glu Ser Arg	5 Ala Gly Lys Pro Pro Gly 189: His Val 1910 Ala Phe	Ser M 1880 Ala L 5 Gly P Ala S	1850 Ser Gly 665 Met Glu Geu Val Pro Gly 1 Ser Arg	Ala Asp Val 19 Glu 915	Val 188 Gly 00 Gly Asn	187 Ile 85 Val Ala	Ala  O  Asn  Ile  Val  193  Thr	Phe Leu Cys Gln 1920 Ala	Leu Ala Trp Pro	Pro Ala Met
1875 Gly Ile Let 1890 Ile Leu Arg 1905 Asn Arg Let	184: Gly 360 Glu Ser Arg Ile 192: Thr	5 Ala Gly Lys Pro Pro Gly 189: His Val 1910 Ala Phe 5 Glu Ser	Ser M 1880 Ala L 5 Gly P Ala S Asp A	1850 Ser Gly 365 Met Glu Geu Val Pro Gly 1 3er Arg 1930 Ala Ser	Ala Asp Val 19 Glu 915 Gly	Leu Val 188 Gly 00 Gly Asn	187 Fle Val Ala His Val	Ala  O  Asn  Ile  Val  193  Thr	Phe Leu Cys Gln 192 Ala 5	Leu Ala Trp Pro	Pro Ala Met Thr
1 Met Ser Gly 1875 Gly Ile Leu 1890 Ile Leu Arg 1905 Asn Arg Leu His Tyr Val	184: Gly 360 Glu Ser Arg Ile 192: Thr	5 Ala Gly Lys Pro Pro Gly 189 His Val 1910 Ala Phe 5 Glu Ser	Ser M 1880 Ala L 5 Gly P Ala S Asp A	1850 Ser Gly 365 Met Glu Geu Val Pro Gly 1 3er Arg 1930 Ala Ser	Ala Asp Val 19 Glu 915 Gly	Leu Val 188 Gly 00 Gly Asn	187 Ile 85 Val Ala His Val 195 Leu	Ala  O  Asn  Ile  Val  193  Thr	Phe Leu Cys Gln 192 Ala 5	Leu Ala Trp Pro	Pro Ala Met Thr

1	970				197	ő			19	80					
Trp	Asp	Trp	Va 1	Сув	Thr	île	Leu	Thr	Asp	Phe	Lys	Asn	Trp	Leu	Thr
1985	5			199	0			1	995				2000	1	
ser	Lys	Leu	Phe	Pro	Lys	Leu	Pro	Gly	Leu	Pro	Phe	Ile	Ser	Сув	Gln
			200	5			2	010				2015	š		
Lys	Gly	Tyr	Lys	Gly	val	Trp	Ala	Gly	Thr	Gly	Ile	Met	Thr	Thr	Arg
		20	20			2	025				203	0			
Сув	Pro	Cys	Gly	Ala	Asn	Ile	Ser	Gly	Asn	Val	Arg	Leu	Gly	Ser	Met
	20	35				2040				20	45				
Arg	Ile	Thr	Gly	Pro	Lys	Thr	Сув	Met	Asn	Thr	Trp	Gln	Gly	Thr	Phe
2	050				205	5			20	60					
Pro	Ile	Asn	Cys	Tyr	Thr	Glu	Gly	Gln	Сув	Ala	Pro	Lys	Pro	Pro	Thr
206	5			207	0			2	075				208	)	
Asn	Tyr	Lys	Thr	Ala	Ile	Trp	Arg	Va1	Ala	Ala	Ser	Glu	Tyr	Ala	Glu
			208	15			2	090				209	5		
Val	Thr	Gln	His	Gly	Ser	Tyr	Ser	Tyr	Val	Thr	Gly	Leu	Thr	Thr	Asp
		21	0 0			2	105				211	.0			
Asn	Leu			Pro	Сув			Pro	Ser	Pro			Phe	Ser	Trp
Asn				Pro	Сув		Leu	Pro	Ser	Pro	Glu		Phe	Ser	Trp
	2	Ьув 115	Ile			Gln 2120	Leu			21	G1 u 25	Phe	Phe Lys		
val	2	Ьув 115	Ile			Gln 2120 His	Leu		Ala	21	G1 u 25	Phe			
val 2	2: Asp 130	Lys 115 Gly	Ile Val	Gln	11e 213	Gln 2120 His 5	Leu Arg	Phe	Ala 2	21: Pro 40	Glu 25 Thr	Phe		Pro	Phe
val 2	Asp 130 Arg	Lys 115 Gly	Ile Val	Gln	Ile 213 Ser	Gln 2120 His 5	Leu Arg	Phe Val	Ala 2	21: Pro 40	Glu 25 Thr	Phe	Lys	Pro Ala	Phe
Val 2 Phe 214	Asp 130 Arg	Lys 115 Gly Asp	Ile Val Glu	Gln Val	11e 213 Ser	Gln 2120 His 5	Leu Arg Cys	Phe Val	Ala 21 Gly 155	Pro .40	Glu 25 Thr Asn	Phe Pro	Lys	Pro Ala	Phe Val
Val 2 Phe 214	Asp 130 Arg	Lys 115 Gly Asp	Ile Val Glu	Gln Val 219 Pro	11e 213 Ser	Gln 2120 His 5	Leu Arg Cys	Phe Val	Ala 21 Gly 155	Pro .40	Glu 25 Thr Asn	Phe Pro	Lys Tyr 216 Val	Pro Ala	Phe Val
Val 2 Phe 214: Gly	Asp 130 Arg 5	Lys 115 Gly Asp Gln	Tle Val Glu Leu 216	Gln Val 219 Pro	11e 213 Ser 0 Cys	Gln 2120 His 5 Phe Glu	Leu Arg Cys Pro	Phe Val 2 Glu 170	Ala 21 Gly 155 Pro	Pro .40 Leu	Glu 25 Thr Asn Ala	Phe Pro Ser Asp	Lys Tyr 216 Val	Pro Ala Leu	Phe Val Arg
Val 2 Phe 214: Gly	Asp 130 Arg 5	Lys 115 Gly Asp Gln Leu	Tle Val Glu Leu 216	Gln Val 219 Pro	11e 213 Ser 0 Cys	Gln 2120 His 5 Phe Glu	Leu Arg Cys Pro	Phe Val 2 Glu 170	Ala 21 Gly 155 Pro	Pro .40 Leu	Glu 25 Thr Asn Ala	Phe Pro Ser Asp 217	Lys Tyr 216 Val	Pro Ala Leu	Phe Val Arg
Val 2 Phe 214: Gly Ser	Asp 130 Arg Ser	Lys 115 Gly Asp Gln Leu 21	Val Glu Leu 216 Thr	Gln Val 219 Pro 5	11e 213 Ser 0 Cys	Gln 2120 His 5 Phe Glu Pro	Leu Arg Cys Pro 2 His	Phe Val 2 Glu 170 Ile	Ala 21 Gly 155 Pro	Pro 40 Leu Asp	Glu 25 Thr Asn Ala Glu 219	Phe Pro Ser Asp 217 Thr	Lys Tyr 216 Val	Pro Ala Leu Ala	Phe Val Arg
Val 2 Phe 214: Gly Ser	Asp 130 Arg Ser Met	Lys 115 Gly Asp Gln Leu 21	Val Glu Leu 216 Thr	Gln Val 219 Pro 5	11e 213 Ser 0 Cys	Gln 2120 His 5 Phe Glu Pro	Leu Arg Cys Pro 2 His	Phe Val 2 Glu 170 Ile	Ala 21 Gly 155 Pro	Pro 40 Leu Asp	Glu 25 Thr Asn Ala Glu 219 Ser	Phe Pro Ser Asp 217 Thr	Lys Tyr 216 Val 5	Pro Ala Leu Ala	Phe Val Arg
val 2 Phe 214! Gly Ser	Asp 130 Arg 5 Ser Met Leu 2	Lys Gly Asp Gln Leu 21 Ala	Val Glu Leu 216 Thr 80 Arg	Gln Val 219 Pro 55 Asp	Ile 213 Ser 60 Cys Pro	Gln 2120 His 5 Phe Glu Pro ::	Leu Arg Cys Pro 2 His	Phe Val 2 Glu 170 Ile	Ala 2: Gly 155 Pro Thr	Pro 40 Leu Asp Ala Ala	Glu 25 Thr Asn Ala Glu 219 Ser	Phe Pro Ser Asp 217 Thr	Lys Tyr 216 Val 5 Ala	Pro Ala Leu Ala	Phe Val Arg
Val 2 Phe 214: Gly Ser Arg	Asp 130 Arg 5 Ser Met Leu 2	Lys Gly Asp Gln Leu 21 Ala	Val Glu Leu 216 Thr 80 Arg	Gln Val 219 Pro 55 Asp	Ile 213 Ser 60 Cys Pro	Gln 2120 His Fhe Glu Pro 2200 Leu	Leu Arg Cys Pro 2 His	Phe Val 2 Glu 170 Ile	Ala 21 Gly 155 Pro Thr	Pro 40 Leu Asp Ala Ala	Glu 25 Thr Asn Ala Glu 219 Ser	Phe Pro Ser Asp 217 Thr	Lys Tyr 216 Val 5 Ala	Pro Ala Leu Ala	Phe Val Arg Arg

2225	2	230			2	235				224	)	
Val Ala Gln	Thr Gl	u Pro	Glu	Ser	Arg	va1	Pro	Val	Leu	qsA	Phe	Leu
	2245			2	250				2255	5		
Glu Pro Met	Ala Gl	u Glu	Glu	ser	Asp	Leu	Glu	Pro	ser	Ile	Pro	Ser
22	60		2	265				227	0			
Glu Cys Met	Leu Pr	o Arg	Ser	Gly	Phe	Pro	Arg	Ala	Leu	Pro	Ala	Trp
2275			2280				22	85				
Ala Arg Pro	Asp Ty	r Asn	Pro	Pro	Leu	Va1	Glu	ser	Trp	Arg	Arg	Pro
2290		229	5			23	00					
Asp Tyr Gln	Pro Pr	o Thr	Val	Ala	Gly	Cys	Ala	Leu	Pro	Pro	Pro	Lys
2305	2	310			2	315				232	)	
Lys Ala Pro	Thr Pr	o Pro	Pro	Arg	Arg	Arg	Arg	Thr	Va1	Gly	Leu	Ser
	2325			2	330				233	5		
Glu Ser Thr	Ile Se	r Glu	Ala	Leu	Gln	Gln	Leu	Ala	Ile	ьув	Thr	Phe
23	4 0		:	345				235	0			
Gly Gln Pro	Pro Se	r Ser	Gly	Asp	Ala	Gly	Ser	Ser	Thr	Gly	Ala	Gly
2355			2360				23	65				
2355 Ala Ala Glu	ser Gl	y Gly			Ser	Pro			Pro	Ala	Pro	Ser
	Ser Gl	y Gly 237	Pro		Ser				Pro	Ala	Pro	ser
Ala Ala Glu		237	Pro 5	Thr		23	Gly 80	Glu				
Ala Ala Glu 2370	Ser Al	237	Pro 5	Thr	Pro	23	Gly 80	Glu			Pro	
Ala Ala Glu 2370 Glu Thr Gly	Ser Al	237 a Ser 390	Pro 5 Ser	Thr Met	Pro 2	23 Pro 395	Gly 180 Leu	Glu Glu	Gly	Glu 240	Pro	Gly
Ala Ala Glu 2370 Glu Thr Gly 2385	Ser Al	237 a Ser 390	Pro 5 Ser	Thr Met Gln	Pro 2	23 Pro 395	Gly 180 Leu	Glu Glu	Gly	Glu 240 Pro	Pro	Gly
Ala Ala Glu 2370 Glu Thr Gly 2385	Ser Al 2 Leu Gl 2405	237 a Ser 390 u Ser	Pro 5 Ser Asp	Thr Met Gln 2	Pro 2 Val 410	Pro 395 Glu	Gly 80 Leu Leu	Glu Glu	Gly Pro 241	Glu 240 Pro	Pro Pro	Gly Gln
Ala Ala Glu 2370 Glu Thr Gly 2385 Asp Pro Asp	Ser Al 2 Leu Gl 2405	237 a Ser 390 u Ser	Pro Ser Asp	Thr Met Gln 2	Pro 2 Val 410	Pro 395 Glu	Gly 80 Leu Leu	Glu Glu	Gly Pro 241: Trp	Glu 240 Pro	Pro Pro	Gly Gln
Ala Ala Glu 2370 Glu Thr Gly 2385 Asp Pro Asp	Ser Al 2 Leu Gl 2405 Val Al	237 a Ser 390 u Ser a Pro	Pro Ser Asp Gly	Thr Met Gln 2 Ser 3425	Pro 2 Val 410 Gly	Pro 395 Glu ser	Gly 80 Leu Leu Gly	Glu Glu Gln Ser 243	Pro 241: Trp	Glu 240 Pro 5	Pro Pro	Gly Gln Cys
Ala Ala Glu 2370 Glu Thr Gly 2385 Asp Pro Asp Gly Gly Gly Gly	Ser Al 2 Leu Gl 2405 Val Al	237 a Ser 390 u Ser a Pro	Pro Ser Asp Gly	Met Gln 2 Ser 2425 Val	Pro 2 Val 410 Gly	Pro 395 Glu ser	Gly 80 Leu Leu Gly	Glu Glu Gln Ser 243 Met	Pro 241: Trp	Glu 240 Pro 5	Pro Pro	Gly Gln Cys
Ala Ala Glu 2370 Glu Thr Gly 2385 Asp Pro Asp Gly Gly Gly 24 Ser Glu Glu	Ser Al  2 Leu Gl 2405 Val Al  220 Asp As	237 a Ser 390 u Ser a Pro	Pro Ser Asp Gly Thr	Met Gln 2 Ser 2425	Pro 2 Val 410 Gly Cys	Pro 395 Glu Ser Cys	Gly 80 Leu Leu Gly Ser 24	Glu Glu Gln Ser 243 Met	Pro 241: Trp 0	Glu 240 Pro Ser	Pro Pro Thr	Gly Gln Cys
Ala Ala Glu 2370 Glu Thr Gly 2385 Asp Pro Asp Gly Gly Gly 24 Ser Glu Glu 2435	Ser Al  2 Leu Gl 2405 Val Al  220 Asp As	237 a Ser 390 u Ser a Pro	Pro Ser Asp Gly Thr 24440 Pro	Met Gln 2 Ser 2425	Pro 2 Val 410 Gly Cys	Pro 395 Glu Ser Cys	Gly 80 Leu Leu Gly Ser 24	Glu Glu Gln Ser 243 Met	Pro 241: Trp 0	Glu 240 Pro Ser	Pro Pro Thr	Gly Gln Cys
Ala Ala Glu 2370 Glu Thr Gly 2385 Asp Pro Asp Gly Gly Gly 24 Ser Glu Glu 2435 Thr Gly Ala	Ser Al  2 Leu Gl 2405 Val Al  20 Asp As	237 a Ser 390 u Ser a Pro p Thr	Pro 5 Ser Asp Gly : Thr 2440 Pro	Thr Met Gln 2 Ser 2425 Val	Pro 2 Val 410 Gly Cys	Pro 395 Glu Ser Cys	Gly Leu Gly Ser 24 Glu	Glu Glu Gln Ser 243 Met 45 Glu	Gly Pro 241: Trp 0 ser	Glu 240 Pro 5 Ser Tyr	Pro Pro Pro Thr	Gly Gln Cys Trp
Ala Ala Glu 2370 Glu Thr Gly 2385 Amp Pro Amp Gly Gly Gly 24 Ser Glu Glu 2435 Thr Gly Ala 2450	Ser Al 2 Leu Gl 2405 Val Al 220 Asp As Leu Il Leu Se	237 a Ser 390 u Ser a Pro p Thr	Pro 5 Ser Asp Gly : Thr 2440 Pro	Thr Met Gln 2 Ser 2425 Val	Pro 2 Val 410 Gly Cys Ser	Pro 395 Glu Ser Cys	Gly Leu Gly Ser 24 Glu	Glu Glu Gln Ser 243 Met 45 Glu	Gly Pro 241: Trp 0 ser	Glu 240 Pro 5 Ser Tyr	Pro Pro Pro Pro Thr Leu Val	Gly Gln Cys Trp

			248	5			2	490				2495	5		
Asp	Arg	Thr	Gln	Val	Leu	Asp	Ala	His	Tyr	Asp	ser	Val	Leu	Lys	Asp
		25	00			2	505				251	0			
11e	Lys	Leu	Ala	Ala	Ser	Lys	Val	Ser	Ala	Arg	Leu	Leu	Thr	Leu	Glu
	25	515				2520				25	25				
G1u	Ala	Сув	Gln	Leu	Thr	Pro	Pro	His	Ser	Ala	Arg	ser	гàз	Tyr	Gly
2	530				253	5			25	40					
Phe	Gly	Ala	Lys	Glu	<b>v</b> al	Arg	Ser	Leu	Ser	Gly	Arg	Ala	Val	Asn	His
2545	5			255	0			2	555				2560	)	
Ile	ьув	Ser	Val	Trp	Lys	Asp	Leu	Leu	Glu	Asp	Pro	Gln	Thr	Pro	Ile
			256	5			2	570				257	5		
Pro	Thr	Thr	Ile	Met	Ala	Lys	Asn	Glu	Va1	Phe	Cys	Val	Asp	Pro	Ala
		25	80			:	2585				259	0			
Lys	Gly	Gly	ьув	Lys	Pro	Ala	Arg	Leu	Ile	Va1	Tyr	Pro	Авр	Leu	Gly
	25	95				2600				26	05				
va1	Arg	Va1	Cys	Glu	ьув	Met	Ala	Leu	Tyr	Asp	Ile	Thr	Gln	Lys	Leu
2	610				261	5			26	20					
		Ala	Val	Met			Ser	туг			Gln	туг	Ser	Pro	Ala
	G1n	Ala	Val	Met 263	Gly		Ser				Gln	Туг	Ser 264		Ala
Pro 262	Gln				Gly	Ala		2	Gly 635	Phe			264	0	
Pro 262	Gln			263 Tyr	Gly	Ala	Ľуз	2	Gly 635	Phe			264 Lys	0	
Pro 262 Gln	Gln Arg	Val	Glu 264	263 Tyr	Gly 10 Leu	Ala	Lys 2	2 Ala 650	Gly 635 Trp	Phe	G1 u	Lys 265	264 Lys 5	Asp	Pro
Pro 262 Gln	Gln Arg	Val	Glu 264	263 Tyr 15	Gly 10 Leu	Ala Leu Thr	Lys 2	2 Ala 650	Gly 635 Trp	Phe	G1 u	Lys 265 Thr	264 Lys 5	Asp	Pro
Pro 2629 Gln Met	Gln Arg Gly	Val	Glu 264 Ser 60	263 Tyr 15	Gly 0 Leu Asp	Ala Leu Thr	Lys 2 Arg 2665	2 Ala 650 Cys	Gly 635 Trp Phe	Phe Ala Asp	Glu Ser 267	Lys 265 Thr	264 Lys 5 Val	Asp Thr	Pro Glu
Pro 2629 Gln Met	Gln Arg Gly	Val	Glu 264 Ser 60	263 Tyr 15 Tyr	Gly 10 Leu Asp Glu	Ala Leu Thr	Lys 2 Arg 2665 Ser	2 Ala 650 Cys	Gly 635 Trp Phe	Phe Ala Asp	Glu Ser 267 Ala	Lys 265 Thr	264 Lys 5 Val	Asp Thr	Pro Glu
Pro 2629 Gln Met	Gln Arg Gly Asp	Val Phe 26 Ile	Glu 266 Ser 60 Arg	263 Tyr 15 Tyr	Gly 10 Leu Asp Glu	Ala Leu Thr Glu 2680	Lys 2 Arg 2665 Ser	Ala 650 Cys	Gly 635 Trp Phe Tyr	Phe Ala Asp Gln 26 Thr	Glu Ser 267 Ala 85	Lys 265 Thr 0 Cys	Lys  Val	Asp Thr	Pro Glu Pro
Pro 2629 Gln Met Arg Glu	Gln Arg Gly Asp Glu 690	Phe 26 Ile 575 Ala	Glu 266 Ser 60 Arg	263 Tyr 15 Tyr Thr	Gly Leu Asp Glu Ala 269	Ala Leu Thr Glu 2680 Ile	Lys 2 Arg 2665 Ser His	Ala 650 Cys Ile Ser	Gly 635 Trp Phe Tyr Leu	Phe Ala Asp Gln 26 Thr	Glu Ser 267 Ala 85 Glu	Lys 265 Thr 0 Cys	Lys  Val  Ser  Leu	Asp Thr Leu	Pro Glu Pro Val
Pro 2629 Gln Met Arg Glu	Gln Arg Gly Asp Glu 690	Phe 26 Ile 575 Ala	Glu 266 Ser 60 Arg	263 Tyr 5 Tyr	Gly Leu Asp Glu Ala 269	Ala Leu Thr Glu 2680 Ile	Lys 2 Arg 2665 Ser His	Ala 650 Cys Ile Ser	Gly 635 Trp Phe Tyr Leu	Phe Ala Asp Gln 26 Thr	Glu Ser 267 Ala 85 Glu	Lys 265 Thr 0 Cys	Leu	Asp Thr Leu Tyr	Pro Glu Pro Val
Pro 2623 Gln Met Arg Glu 21 Gly 270	Gln Arg Gly Asp 20 Glu Glu Gly Gly	Val Phe 26 Ile 675 Ala	Glu 266 Ser 60 Arg Arg	263 Tyr 45 Tyr Thr Thr Phe	Gly  Leu  Asp  Glu  Ala  269  Asn	Ala Leu Thr Glu 2680 Ile 5	Lys 2 Arg 2665 Ser His	Ala 650 Cys Ile Ser Gly	Gly 635 Trp Phe Tyr Leu 2: Gln 715	Phe Ala Asp Gln 26 Thr	Glu Ser 267 Ala 85 Glu Cys	Lys 265 Thr 0 Cys Arg	Leu Tyr	Asp Thr Leu Tyr Arg	Pro Glu Pro Val
Pro 2623 Gln Met Arg Glu 21 Gly 270	Gln Arg Gly Asp 20 Glu Glu Gly Gly	Val Phe 26 Ile 675 Ala	Glu 266 Ser 60 Arg Arg	263 Tyr Thr Thr Phe 273 Gly	Gly Leu Asp Glu Ala 269 Asn	Ala Leu Thr Glu 2680 Ile 5	Lys 2 Arg 22665 Ser His	Ala 650 Cys Ile Ser Gly 2	Gly 635 Trp Phe Tyr Leu 2: Gln 715	Phe Ala Asp Gln 26 Thr	Glu Ser 267 Ala 85 Glu Cys	Lys 265 Thr 0 Cys Arg Gly	Leu Tyr 272	Asp Thr Leu Tyr Arg	Pro Glu Pro Val
Pro 2622 Gln Met Arg Glu 2 Gly Cys	Gln  Arg  Gly  Asp  2  Glu  690  Gly  Arg	Phe 26 11e 575 Ala	Glu 264 Ser 60 Arg Met Ser 272	263 Tyr Thr Thr Phe 273 Gly	Gly Leu Asp Glu Ala 269 Asn O Val	Ala Leu Thr : Glu 2680 Ile 5 Ser	Lys 2 Arg 2665 Ser His Lys	Ala 650 Cys Ile Ser Gly 2 Thr 730	Gly 635 Trp Phe Tyr Leu 2' Gln 715	Phe Ala Asp Gln 26 Thr 700 Thr	Glu Ser 267 Ala 85 Glu Cys	Lys 265 Thr 0 Cys Arg Gly Asn 273	264' Lys  Val  Ser  Leu  Tyr  272 Thr	Thr Leu Tyr Arg	Pro Glu Pro Val Arg

	27	40			2	745				275	0			
Pro Thr	Met	Leu	Val	Cys	Gly	Asp	Asp	Leu	Val	Val	Ile	ser	Glu	ser
27	755				2760				27	65				
Gln Gly	Thr	Glu	Glu	Авр	Glu	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala
2770				277	5			27	80					
Met Thr	Arg	Tyr	ser	Ala	pro	pro	Gly	Asp	Pro	Pro	Arg	Pro	Glu	Tyr
2785			279	0			2	795				2800	)	
Asp Leu	Glu	Leu	Île	Thr	ser	Cys	ser	ser	Asn	Val	Ser	va1	Ala	Leu
		280	15			2	810				281	5		
Gly Pro	Arg	Gly	Arg	Arg	Arg	Tyr	Tyr	Leu	Thr	Arg	Asp	Pro	Thr	Thr
	28	20			2	825				283	0			
Pro Leu	Ala	Arg	Ala	Ala	Trp	Glu	Thr	Val	Arg	His	Ser	Pro	Ile	Asn
28	3 5				2840				28	45				
Ser Trp	Leu	Gly	Asn	Ile	Ile	Gln	Tyr	Ala	Pro	Thr	Ile	Trp	Val	Arg
2850				285	5			28	860					
Met Val	Leu	Met	Thr	His	Phe	Phe	ser	Ile	Leu	Met	Val	Gln	Asp	Thr
2865			287	0			2	875				288	0	
2865 Leu Asp	Gln	Asn			Phe	Glu			Gly	ser	Val			Val
	Gln	Asn 288	Leu		Phe				Gly	Ser	Val 289	Tyr		Val
		288	Leu 5	Asn		2	Met 890	Tyr			289	Tyr 5	Ser	
Leu Asp	Leu	288	Leu 5	Asn	Ala	2	Met 890	Tyr			289 His	Tyr 5	Ser	
Leu Asp	Leu 29	288 Asp 00	Leu 5 Leu	Asn	Ala	2 Ile 1905	Met 890 Ile	Tyr Glu	Arg	Leu 291	0 1139	Tyr 5 Gly	Ser Leu	Asp
Leu Asp Asn Pro	Leu 29	288 Asp 00	Leu 5 Leu	Asn	Ala	2 Ile 905 Ser	Met 890 Ile	Tyr Glu	Arg	Leu 291 Leu	0 1139	Tyr 5 Gly	Ser Leu	Asp
Leu Asp Asn Pro	Leu 29 Ser	288 Asp 00 Met	Leu 5 Leu His	Asn Pro	Ala Tyr 2920	Ile 1905 Ser	Met 890 Ile His	Tyr Glu His	Arg Glu 29	Leu 291 Leu 25	289 His O Thr	Tyr 5 Gly Arg	Ser Leu Val	Asp Ala
Leu Asp Asn Pro Ala Phe	Leu 29 Ser	288 Asp 00 Met	Leu 5 Leu His	Asn Pro	Ala Tyr 2920 Gly	Ile 1905 Ser	Met 890 Ile His	Tyr Glu His	Arg Glu 29	Leu 291 Leu 25	289 His O Thr	Tyr 5 Gly Arg	Ser Leu Val	Asp Ala
Leu Asp Asn Pro Ala Phe 23 Ser Ala	Leu 29 Ser 915 Leu	Asp 00 Met Arg	Leu 5 Leu His	Asn Pro Thr Leu 293	Ala Tyr 2920 Gly 5	Ile 1905 Ser	Met 890 Ile His	Glu His Pro	Arg Glu 29 Leu	Leu 291 Leu 25 Arg	289 His O Thr	Tyr 5 Gly Arg Trp	Ser Leu Val	Asp Ala Ser
Leu Asp Asn Pro Ala Phe 29 Ser Ala 2930	Leu 29 Ser 915 Leu	Asp 00 Met Arg	Leu 5 Leu His	Asn Pro Thr Leu 293 Arg	Ala Tyr 2920 Gly 5	Ile 1905 Ser	Met 890 Ile His Pro	Glu His Pro	Arg Glu 29 Leu	Leu 291 Leu 25 Arg	289 His O Thr	Tyr 5 Gly Arg Trp	Ser Leu Val Lys	Asp Ala Ser
Leu Asp Asn Pro Ala Phe 29 Ser Ala 2930 Arg Ala	Leu 29 Ser 15 Leu Arg	Asp 00 Met Arg	Leu E5 Leu His Lys Val	Asn Pro Thr Leu 293 Arg	Ala Tyr 2920 Gly 5 Ala	Ile 1905 Ser Ala Ser	Met 890 Ile His Pro Leu 2	Glu His Pro 25 Tle 955	Glu 29 Leu 340 Ser	Leu 291 Leu 25 Arg	289 His O Thr Val	Tyr  Gly  Arg  Trp  Gly  296	Leu Val Lys	Asp Ala Ser
Leu Asp Asn Pro Ala Phe 29 Ser Ala 2930 Arg Ala 2945	Leu 29 Ser 15 Leu Arg	Asp 00 Met Arg	Leu His Lys Val 295	Asn Pro Thr Leu 293 Arg	Ala Tyr 2920 Gly 5 Ala	Ile 1905 Ser Ala Ser	Met 890 Ile His Pro Leu 2	Glu His Pro 25 Tle 955	Glu 29 Leu 340 Ser	Leu 291 Leu 25 Arg	289 His O Thr Val	Tyr  Gly  Gly  Trp  Gly  2960	Leu Val Lys	Asp Ala Ser
Leu Asp Asn Pro Ala Phe 29 Ser Ala 2930 Arg Ala 2945	Leu 29 Ser 15 Leu Arg	Asp 00 Met Arg Ala Gly	Leu His Lys Val 298 Arg	Asn Pro Thr Leu 293 Arg 50	Ala Tyr 2920 Gly 5 Ala	Ile 1905 Ser Ala Ser	Met 890 Ile His Pro Leu 2 Asn	Tyr Glu His Pro 29 Fle 955	Arg Glu 29 Leu 940 Ser	Leu 291 Leu 25 Arg Arg	289 His O Thr Val Gly Lys 297	Tyr  5  Gly  Arg  Trp  Gly  2960  Thr	Ser Leu Val Lys Lys	Asp Ala Ser Ala Leu
Asn Pro Ala Phe 29 Ser Ala 2930 Arg Ala 2945 Ala Val	Leu 29 Ser 15 Leu Arg	Asp 00 Met Arg Ala Gly 296 Pro	Leu His Lys Val 298 Arg	Asn Pro Thr Leu 293 Arg 50	Ala Tyr 2920 Gly Ala Leu Glu	Ile 1905 Ser Ala Ser	Met 890 Ile His Pro Leu 2 Asn	Tyr Glu His Pro 29 Fle 955	Arg Glu 29 Leu 940 Ser	Leu 291 Leu 25 Arg Arg	289 His O Thr Val Gly Lys 297 Leu	Tyr  5  Gly  Arg  Trp  Gly  2960  Thr	Ser Leu Val Lys Lys	Asp Ala Ser Ala Leu

3005 3000 2995 Ala Arg Pro Arg Ser Leu Leu Phe Gly Leu Leu Leu Phe Val Gly 3020 3010 3015 Val Gly Leu Phe Leu Leu Pro Ala Arg 3030 3025 <210> 5 <211> 9674 <212> DNA <213> Hepatitis C virus <220> <221> CDS <222> (341)..(9442) <400> 5 accegecet aataggggeg acacteegee atgaateact eccetgtgag gaactactgt 60 cttcacgcag aaagcgtcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120 ccccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg 180 aagactgggt cetttettgg ataaacccac tetatgeeeg gecatttggg egtgeeeeg 240 caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300 tgettgegag tgecceggga ggtetegtag acegtgeace atg age aca aat eee 355 Met Ser Thr Asn Pro 1 5 aaa cet caa aga aaa ace aaa aga aac act aac egt ege cea caa gac 403

Lys	Pro	Gln	Arg	Lys	Thr	Lys	Arg	Asn	Thr	Asn	Arg	Arg	Pro	Gln	Asp	
			1	.0				15				20				
gtt	aag	ttt	ccg	ggc	ggc	ggc	cag	atc	gtt	ggc	gga	gta	tac	ttg	ttg	451
Val	Гув	Phe	Pro	Gly	Gly	Gly	Gln	Ile	Val	Gly	Gly	Val	Tyr	Leu	Leu	
			25				3 0				3 9	;				
ccg	cgc	agg	ggc	ccc	agg	ttg	ggt	gtg	cgc	gcg	aca	agg	aag	get	tcg	499
Pro	Arg	Arg	gly	Pro	Arg	Leu	Gly	Va1	Arg	λla	Thr	Arg	Lys	Ala	ser	
		40				45				5	0					
gag	agg	tee	cag	cca	cgt	ggg	agg	cgc	cag	ccc	atc	ccc	aaa	cat	cgg	547
Glu	Arg	Ser	Gln	Pro	Arg	Gly	Arg	Arg	Gln	Pro	Ile	Pro	гЛа	His	Arg	
	55				60				•	5						
aga	tee	act	ggc	aag	tcc	tgg	ggg	aag	cca	gga	tac	aac	tgg	aaa	ctg	595
Arg	Ser	Thr	Gly	Lys	Ser	Trp	Gly	Lys	Pro	Gly	Tyr	Pro	Trp	Pro	Leu	
70				7	5				80				85			
tat	aaa	aat	gag	999	ctc	ggt	tgg	gca	gga	tgg	ctc	ctg	tcc	cct	cga	643
Tyr	Gly	Asn		Gly	Leu	Gly	Trp		Gly	Trp	Leu		Ser	Pro	λrg	
			9	0				95				100				
				tca												691
Gly	Ser			Ser	Trp	Gly		Asn	Asp	Pro	Arg	His	Arg	Ser	Arg	
		11	05				110				11!	5				
				gtc												739
Asn			Lys	Va1	Ile		Thr	Leu	Thr			Phe	Ala	Asp	Leu	
	1	20				125				1, 3	0					
ttg	99 <b>9</b>	tac	gtc	ccc	gtc	gta	ggc	gcc	ccg	ctt	agt	ggc	gtt	gcc	agt	787

Leu	Gly	Tyr	Val	Pro	Va1	val	Gly	Ala	Pro	Leu	Ser	Gly	Val	Ala	Ser	
	135				140				1	45						
gct	ctc	gcg	cac	ggc	gtg	aga	gtc	ctg	gag	gac	999	gtt	aat	ttt	gca	835
Ala	Leu	Ala	His	Gly	Val	Arg	val	Leu	Glu	Asp	Gly	Val	Asn	Phe	Ala	
150				15	5			3	160				165			
aca	ggg	aac	tta	cct	ggt	tgc	tcc	ttt	tet	atc	ttc	ttg	ctg	gcc	cta	883
Thr	Gly	Asn	Leu	Pro	Gly	Сув	ser	Phe	ser	Ile	Phe	Leu	Leu	Ala	Leu	
			17	0				175				180	1			
ctg	tcc	tgc	atc	act	act	çeg	gtc	tet	gct	gtc	caa	gtg	aag	aac	acc	931
Leu	Ser	Cys	Ile	Thr	Thr	Pro	٧al	Ser	Ala	Val	Gln	Val	Lys	Asn	Thr	
		1	85				190				19	5				
agc	aac	gcc	tat	atg	gcg	act	aac	gac	tgt	tcc	aat	gac	ago	ato	act	979
ser	Asn	Ala	Tyr	Met	Ala	Thr	Asn	Asp	Сув	ser	Asn	Asp	Sex	110	Thr	
		200				205	5			2	10					
tgg	cag	ctt	gag	gcc	gca	gto	cto	cat	gto	ccc	ggg	tgo	gto	009	g tgc	1027
															Cys	
-	215				22					225						
qaq	aaa	atg	999	aac	aca	tca	cgg	, tgc	tgg	ata	a cca	gto	to	a cc	a aac	1075
Glu	Ly	. Met	: Gly	/ Asr	Thr	Sei	Arg	Cys	Tr	110	e Pro	val	L Se	r Pr	o Asn	
230					35				240				24			
gts	ge	gt:	gegg	g çaş	cet	gge	ge	cto	ac	g cg	9 99	c tt	g cg	g ac	g cac	1123
															r His	
				50				255				26				
ate	ga	c at	g gto	gt	gtt	g to	e ge	c ac	g ct	c tg	c tc	c gc	t ct	c ta	c gtg	1171

Ile	Asp	Met	Val	Va1	Leu	Ser	Ala	Thr	Leu	Сув	Ser	Ala	Leu	Tyr	Val	
		2	6 5				270				27	5				
ggg	gac	atc	tgt	ggc	999	gtg	atg	ctc	gcg	tcc	cag	atg	ttc	att	gto	1219
Gly	Asp	Leu	Сув	Gly	Gly	Va1	Met	Leu	Ala	ser	G1n	Met	Phe	11e	Val	
	2	80				285				29	0					
teg	ceg	cag	cac	cac	tgg	ttc	gtg	cag	gaa	tgc	aat	tgc	tcc	atc	tac	1267
ser	Pro	Gln	His	His	Trp	Phe	٧al	Gln	Glu	Сув	Asn	Cys	ser	Ile	Tyr	
	295				300	)			3	05						
cct	ggc	gcc	atc	act	aaa	cac	cgt	atg	gca	tgg	gac	atg	atg	atg	aac	1315
Pro	Gly	Ala	11e	Thr	G1y	His	Arg	Met	Ala	Trp	Asp	Met	Met	Met	Asn	
310				31	5			:	320				325			
tgg	teg	ccc	acg	acc	acc	atg	atc	ctg	geg	tac	gtg	atg	cgc	gtt	acc	1363
Trp	Ser	Pro	Thr	Thr	Thr	Met	Ile	Leu	λla	Tyr	Val	Met	Arg	Val	Pro	
			3 3	0				335				340	)			
gag	atc	atc	ata	qac	atc	att	agc	gga	get	cac	tgg	ggc	gtc	atg	ttt	1411
														Met		
	-		45	-			350				35					
		,														
aac	eta	acc	tac	tte	tet	atq	caq	qga	geg	tgg	geg	aag	gte	gtt	gtc	1459
	-	-												Va 1		
017		360	-1-			365		,			10	•				
		360				303				-						
																1507
		-													999	150,
Ile	Leu	Leu	Leu	Ala	Ser	Gly	Val	Asp			Thr	Thr	Thr	Thr	GIA	
	375				38	0			3	85						
agc	gct	gct	99 <b>9</b>	aga	act	acc	agt	agc	ctg	gaa	agc	gcc	tto	tac	cct	1555

ser	Ala	Ala	Gly	Arg	Thr	Thr	Ser	ser	Leu	Ala	Ser	Ala	Phe	Ser	Pro	
390				39	5			•	00				405			
ggc	gct	cgg	cag	aac	att	cag	ctc	att	aat	acc	aat	ggt	agc	tgg	cac	1603
Gly	Ala	Arg	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr	Asn	Gly	Ser	Trp	His	
			41	.0			•	115				420				
															ttc	1651
Ile	Asn		Thr	Ala	Leu	Asn		Asn	Asp	Ser			Thr	Gly	Phe	
		4	25				430				43	5				
																1699
	-	-	ctg													1033
Phe			Leu	Pne	Tyr	445	нів	гуа	Pne	AB11		ser	GIY	Сув	PIO	
	9	40				445				413						
~=~	cac	cta	tee	acc	tat	cac	aac	atc	aaa	gac	tte	caa	ata	ασa	tgg	1747
	-		Ser													
	455				460	-				65				•	-	
ggc	gcc	ctg	caa	tac	gac	gac	aat	gtc	acc	aat	cca	gaa	gat	atg	agg	1795
Gly	Ala	Leu	Gln	Tyr	Asp	Asp	Asn	Val	Thr	Asn	Pro	Glu	Asp	Met	Arg	
470				47	5				480				485			
cca	tat	tgc	tgg	cac	tac	cca	cca	aaa	cag	tgt	ggc	gta	gtc	ccc	gca	1843
Pro	Tyr	Cys	Trp	His	Tyr	Pro	Pro	Lys	Gln	Сув	Gly	Val	Val	Pro	Ala	
			4 9	0				495				500				
999	acc	gtg	tgc	ggc	cca	gtg	tac	tgt	ttc	acc	cct	agc	ccg	gtg	gta	1891
Gly	Thr	Val	Сув	Gly	Pro	Va1	Tyr	Cys	Phe	Thr	Pro	Ser	Pro	Val	Val	
		5	05				510				51	5				
ata	aac	acq	acc	gat	aga	ctt	aga	ata	cct	act	tac	acq	taa	qqa	gag	1939

Val	Gly	Thr	$\operatorname{Th} r$	Asp	Arg	Leu	Gly	Val	Pro	Thr	Tyr	Thr	Trp	Gly	Glu	
	5	20				525				53	0					
aat	gag	aca	gat	gtc	ttc	cta	ttg	aac	agc	acc	cga	cca	ccg	teg	ggg	1987
Asn	G1u	Thr	Asp	va1	Phe	Leu	Leu	Asn	Ser	Thr	Arg	Pro	Pro	Ser	Gly	
	535				540	)			5	4 5						
tca	tgg	ttt	ggc	tgc	acg	tgg	atg	aac	tcc	act	ggc	ttc	acc	aag	acc	2035
Ser	Trp	Phe	Gly	Сув	Thr	Trp	Met	Asn	Ser	Thr	Gly	Phe	Thr	Lys	Thr	
550				55	5				560				565			
tgc	ggc	gca	cca	ccc	tgc	cgc	act	aga	gct	gac	ttc	aat	acc	agc	aca	2083
Cys	Gly	Ala	Pro	Pro	Сув	Arg	Thr	Arg	Ala	Asp	Phe	Asn	Thr	Ser	Thr	
			57	0				575				580				
gat	ctg	ttg	tgc	ccc	acg	gac	tgt	ttt	aga	aaa	cat	cct	gaa	gcc	act	2131
Asp	Leu	Leu	Сув	Pro	Thr	Азр	Cys	Phe	Arg	Lys	His	Pro	Glu	Ala	Thr	
		5	85				590				59	5				
tac	atc	aaa	tgt	ggt	tcc	999	cct	tgg	ctc	acg	cca	aag	tgt	ctg	gtt	2179
Tyr	Ile	Lys	Cys	Gly	Ser	Gly	Pro	Trp	Leu	Thr	Pro	Lys	Cys	Leu	Val	
		500				605				6	1.0					
gac	tac	ccc	tac	agg	ctc	tgg	cat	tac	cct	tgc	aca	gtc	aat	tac	tcc	2227
Asp	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Сув	Thr	Val	Asn	Tyr	Ser	
	615				62	0			6	25						
acc	ttc	aag	atc	agg	atg	tat	gtg	aaa	gga	gtt	gag	cac	agg	ctc	atg	2275
Thr	Phe	Lys	11e	Arg	Met	Tyr	Va1	Gly	Gly	Val	Glu	His	Arg	Leu	Met	
630				63	15				640				645	5		
nce	aca	tac	aat	tte	act	cat.	aaa	gat	cac	tac	aac	ttq	qaq	qat	agg	2323

Ala	Ala	Суа	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Сув	Asn	Leu	Glu	Asp	Arg	
			65	0				555				660				
gac	aga	agt	caa	cag	act	cct	ctg	ttg	cac	tcc	acc	acg	gaa	tgg	gcc	2371
Asp	Arg	ser	Gln	Gln	Thr	Pro	Leu	Leu	нів	Ser	Thr	Thr	Glu	Trp	Ala	
		6	65				670				67	õ				
att	ttg	ccc	tgc	tet	ttc	tca	gac	ttg	ccc	gct	ttg	tcg	act	ggt	ctt	2419
Ile	Leu	Pro	Сув	Ser	Phe	ser	Asp	Leu	Pro	Ala	Leu	Ser	Thr	Gly	Leu	
	6	680				685				69	0					
ctc	cac	ctc	cac	caa	aat	atc	gtg	gac	gta	caa	tat	atg	tat	ggc	ctg	2467
Leu	His	Leu	His	Gln	Asn	Ile	Va 1	Авр	Val	Gln	Tyr	Met	Tyr	Gly	Leu	
	695				700	)			7	05						
tca	cct	gcc	ctc	aca	caa	tat	atc	gtt	cga	tgg	gag	tgg	gta	gta	ctc	2515
ser	Pro	Ala	Leu	Thr	Gln	Tyr	Ile	Val	Arg	Trp	Glu	Trp	Val	Val	Leu	
710				71	5				720				725			
tta	ttc	ctg	ctc	cta	gcg	gac	gcc	agg	gtc	tgc	gcc	tgc	ttg	tgg	atg	2563
Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Val	Сув	Ala	Cys	Leu	Trp	Met	
			73	0				735				740	)			
ctc	atc	ttg	ctg	ggc	caa	gcc	gaa	gca	gca	ctg	gag	aag	ctg	gtc	gtc	2611
Leu	Ilė	Leu	Leu	Gly	Gln	Ala	Glu	Ala	Ala	Leu	Glu	Гув	Leu	Val	Val	
		7	45				750				75	5				
ttg	cac	gct	gcg	agc	gca	get	age	tgc	aat	ggc	ttc	ctg	tat	ttt	gtc	2659
Leu	His	Ala	Ala	ser	Ala	Ala	ser	Cys	Asn	Gly	Phe	Leu	Tyr	Phe	Val	
		760				765				77	0					
atc	ttt	ctc	gtg	gct	gct	tgg	cac	atc	aag	ggt	agg	gtg	gtc	ccc	ttg	2707

Ile	Phe	Leu	Val	Ala	Ala	Trp	His	11e	Lys	Gly	Arg	Val	Val	Pro	Leu	
	775				780				7	85						
gct	gct	tat	tcc	ctt	act	ggc	ctg	tgg	ccg	ttc	tgc	cta	ctg	ctc	cta	2755
Ala	Ala	Tyr	Ser	Leu	Thr	Gly	Leu	Trp	Pro	Phe	Cys	Leu	Leu	Leu	Leu	
790				79	5			8	300				805			
-	_										tet					2803
Ala	Leu	Pro	Gln	Gln	Ala	туг	Ala	Tyr	Asp	Ala	Ser	Val	His	Gly	Gln	
			81	.0				815				820	1			
															ccg	2851
Va 1	Gly	A1a	Ala	Leu	Leu	Val		Ile	Thr	Leu	Phe		Leu	Thr	Pro	
		8	25				830				83	5				
																2899
															ata	2899
Gly			Thr	Leu	Leu			Ser	Leu			Leu	Cys	Tyr	Leu	
	,	840				845	•			8	5 0					
																2947
															cag	2347
Leu		Leu	Ala	Glu			Va1	Gin			Ala	Pro	ser	Mec	Gln	
	855				86	0			•	865						
												200	ata	+++	tga	2995
															сув	
		GIA	GIY		75	GIY	, 116		880	, Ale	, ALG		88		-,-	
870				۰	75											
					~~				r tac	a ete	tt.	gee	ratio	rett	<b>. g</b> gg	3043
															Gly	
51.0	OIY	val		90				895	1			90				
			٥	- •												
ant	aat	l-ac	cte	, eta	aga	aat	: act	: ttc	ace	1 646	e ato	cca	ı tat	tto	gte	3091

Pro Gly Tyr	Leu Leu Arg	Gly Ala	Leu Thr	Arg Val	Pro Tyr	Phe Val	
90	5	910		915			
aga gcc cac	get etg etg	aga atg	tgc act	atg gtg	agg cac	ete geg	3139
Arg Ala His	Ala Leu Leu	Arg Met	Cys Thr	Met Val	Arg His	Leu Ala	
920		925		930			
ggg ggt agg	tac gtc cas	atg gcg	cta tta	gcc ctt	ggc agg	tgg act	3187
Gly Gly Arg	Tyr Val Gl	Met Ala	Leu Leu	Ala Leu	Gly Arg	Trp Thr	
935	94	0	9	15			
ggc act tac							3235
Gly Thr Tyr		His Leu		Met Ser		Ala Ala	
950	955		960		965		
age gge etg							3283
Ser Gly Leu				GIU Pro	980	Phe Ser	
	970	9	975		980		
ccg atg gag				~~~ ~~~	ann ant	ana tan	3331
Pro Met Glu							3331
98	-	990	,,	995		,-	
,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	•						
ggg gac atc	ttq cac qq	ctt ccc	gtg tcc	gec ega	ctc ggt	Cgg gag	3379
Gly Asp Ile							
1000		1005		1010			
atc ctc ctt	ggc cca gct	gat ggc	tac acc	tcc aag	ggg tgg	aag ctt	3427
Ile Leu Leu	Gly Pro Ala	Asp Gly	Tyr Thr	Ser Lys	Gly Trp	Lys Leu	
1015	10	20	10	25			
cte gcc cec	atc acc gc	tac gcc	cag cag	aca cga	ggt ctc	ttg ggc	3475

Leu	Ala	Pr	0 ]	1e	Thr	Ala	Tyr	Ala	Gln	Gln	Thr	Arg	Gly	Leu	Leu	G.	lу	
1030					103	5			1	040				104	5			
tet	ata	gt	g	gtg	agc	atg	acg	ggg	cgt	gac	aag	aca	gaa	cag	gcc	g	gg	3523
Ser	Ile	Va	.1 '	Val	Ser	Met	Thr	Gly	Arg	qaA	ьув	Thr	Glu	Gln	Ala	G	ly	
				105					055				106					
gag	gto	ce	aa	gtc	ctg	tee	aca	gtc	act	cag	tcc	ttc	ctc	gga	aca	ı t	cc	3571
Glu	Va]	G	ln	۷al	Leu	Ser	Thr	Val	Thr	Gln	ser	Phe	Leu	G1y	Thi	S	er	
			106	5				1070				10	75					
att	tes	1 9	gg	gte	tta	tgg	act	gtt	tac	cac	gga	get	ggc	aac	aa	gē	ıca	3619
Tle	Se:	c G	lv	val	Leu	Trp	Thr	Val	Туг	нів	Gly	Ala	Gly	/ Ası	ı Ly	s :	fhr	
		108					108					090						
	~~		ac	tco	cas	ggc	ccc	qto	acc	cag	ate	g ta	c to	gag	c gc	c	gag	3667
ten	, gc	- 6	o v	Ser	Arc	Gly	Pro	va:	Th	Glr	Me!	t Ty	r Se	r se	r Al	a	Glu	
	1095			,,,,		11					105							
	109	,																
				~+-		g tg		n ag	e ee	: cci	. gg	g ac	c aa	a to	t tt	g	gag	3715
999	ga.	c t	.cg	911	99	y Tr	, Dr	o Se	r Pr	o Pro	o Gl	y Th	r Ly	s Se	r Le	u	Glu	
		рг	ieu	va.		, 11. 115	,			1120					25			
111	10				1.	115												
						a gc			c ct	a ta	t tt	a at	c ac	q cg	ıg a	ac	gct	3763
ceg	g to	t a	acg	cg	. gg	a gc y Al		. 94		9 U.	r Le	n Va	1 Th	r A	rq A	gn	Ala	
Pro	o Cy	s :	rhr			у АТ	a va	I WB	1139					40	-			
				1:	130				113:	,								
													~-		-a -	t a	ete	3811
ga	t g	c	atc	cc	g gc	t cg	a ag	a cg	ic gg	g ga	.c aa	ig cg	19 99	ja y			Lou	
As	p V	1	Ile	Pr	o Al	a Ar	g Ar			у Ая	p Ly			LYA	ra n	-a	Lou	
			1	145				115	0			3	155					
to	e c	cg	aga	ı cc	e e1	t to	g ac	ec ti	g as	ıg gg	g t	ee t	eg g	gg <b>g</b>	ga c	ct	gtg	3859

Ser	Pro	Arg	Pro	Leu	Ser	Thr	Leu	Lys	Gly	Ser	Ser	Gly	Gly	Pro	٧a	ıl	
	11	60				1165				11	70						
att	tgc	ect	agg	ggc	cac	gct	gtc	gga	atc	ttc	cgg	gca	gct	gtg	te	gc	3907
Leu	Cys	Pro	Arg	Gly	His	Ala	Val	Gly	11e	Phe	Arg	Ala	Ala	Va1	C?	ув	
1:	175				118	0			11	85							
tat	cgg	ggt	gtg	gct	aag	tcc	ata	gat	ttc	atc	ccc	gtt	gag	acg	· c	tc	3955
Ser	Arg	G1y	Va1	Ala	Lys	Ser	Ile	Asp	Phe	11e	Pro	Va1			L	eu	
1190	)			11	95			1	200				120	5			
gac	atc	gtc	acg	cgg	tet	ccc	acc	ttt	agt	gac	aac	ago	aca	dCa	ı c	ca	4003
Asp	Ile	Val	Thr	Arg	ser	Pro	Thr	Phe	Ser	Asp	Asn			Pro	) P	ro	
			12	10			:	1215				122	0				
gct	gtg	ccc	gag	acc	tat	cag	gte	999	tac	tts	g cac	gco	ccc	acı	t g	gge	4051
Ala	Va 1	Pro	Glr	Thi	Туз	Glr	Va]	G1y	Tyr	Let			Pro	Th	re	31y	
		3.2	225				1230	)			12	35					
agt	gga	aaa	ago	aco	aaq	gto	000	gto	geç	, ta	c gc	e ge	c cas	9 99	g t	tat	4099
ser	Gly	Lys	se	r Thi	r Ly:	∌ Va:	Pr	o Val	L Ala			a Al	a Gli	n GI	у:	ryr	
	1	240				124	5			1	250						
																	4147
aaa	gts	ct	ggt	g at	c aa	t cc	e te	g gt	g gci	gc	c ac	c ct	g <b>g</b> g	a tt	.t	999	4147
Lys	Va]	Le	ı Va	l Le	u As	n Pr	o Se	r Va				r Le	u GI	ург	ie i	GIĀ	
	1255				12	60				1265							
																	4105
gcg	j ta	e tt	g tc	c aa	g gc	a ca	t gg	c at	c aa	c cc	c aa	c at	t ag	g ac	jt.	gga	4195
Ala	ту	r Le	u Se	r Ly	s Al	a Hi	s Gl	y II			o As	n 11			ır	GIY	
12	10			1	275				1280	)			12	85			
ati	n aq	a ac	t gt	gac	gac	c gg	g ga	ıg cc	c at	t a	ca ta	ic to	ec ac	g t	at	ggt	4243
50	5		-														

Val Arg Thr V	al Thr Thr	Gly Glu Pro	Ile Thr Tyr	Ser Thr Tyr Gly
	1290	1295		1300
				tat gac atc atc 4291
			Gly Gly Ala	Tyr Asp Ile Ile
130	5	1310	131	5
ata tqc qat q	gaa tgc cac	tct gtg gat	gct acc act	att ctc ggc atc 4339
				Ile Leu Gly Ile
1320		1325	1330	
ggg aca gtc	ctt gac caa	gca gag aca	gcc ggg gtc	agg cta act gta 4387
Gly Thr Val I	Leu Asp Gln	Ala Glu Thr	Ala Gly Val	Arg Leu Thr Val
1335	1340	0	1345	
				con cat coc ast 4435
				ccc cat ccc aat 4435 Pro His Pro Asn
Leu Ala Thr	1355		360	1365
1350	1335	•	300	
ata gag gag	qta qcc ctc	gga cag gag	ggt gag atc	ccc ttc tat ggg 4483
				Pro Phe Tyr Gly
	1370	1375		1380
				cac ttg att ttc 4531
Arg Ala Phe	Pro Leu Ser	Tyr Ile Lys		His Leu Ile Phe
138	15	1390	135	95
				4570
				gcc ctt cgg ggc 4579
Cys His Ser	гуз гуз гуз	1405	1410	Ala Leu Arg Gly
1400				
atg ggc ttg	aac get gtg	gca tat tac	aga ggg ttg	gac gtc tcc ata 4627

Met G	ly	Leu	Asn	Ala	Va 1	Ala	Tyr	Tyr	Arg	Gly	Leu	Asp	Val	ser	Ile	
14	15				1420	)			14	25						
ata c																4675
Ile P	ro	Thr	Gln	Gly	Asp	Val	Val			Ala	Thr	Asp			Met	
1430				143	15			1	440				1445	5		
																4723
acg g																4/23
Thr G	ly	Tyr			Asp	Phe			Val	Ile	Asp	146		vaı	Ата	
			145	0			1	455				1464	,			
gtc a								ata	~~~	000	acc	ttc	act	ata	acc	4771
gtc a																
var	ını	14		vai	nop		1470	200			141					
		1.1	0.5													
aca (	na cr	act	atc	cca	caa	qac	qct	gte	tca	cgt	agt	cag	cgc	cga	999	4819
Thr C																
		80				1485					90					
cgc 8	acg	ggt	aga	gga	aga	ctg	ggc	att	tat	agg	tat	gtt	tcc	act	ggt	4867
Arg :	rhr	Gly	Arg	Gly	Arg	Leu	Gly	Ile	Tyr	Arg	Tyr	Va1	ser	Thr	Gly	
14	95				150	0			1	505						
															tac	4915
Glu i	Arg	Λla	Ser	Gly	Met	Phe	Asp	Ser	Val	Va1	Leu	Cys			Tyr	
1510				15	15			1	520				152	5		
															gtc	4963
Asp :	Ala	Gly			Trp	Tyr			ser	Pro	Val			Thi	val	
			15	30				1535				154	U			
												ato				5011
agg -	ctc	agg	gcg	tat	ttc	aac	acc	000	ggc	Lug	,	. grg	cyc	. cag	gac	3011

Arg Leu Arg A	la Tyr Phe	Asn Thr E	ro Gly L	eu Pro Val	Cys Gln	Asp
1545	5	1550		1555		
cac ett gag t	tt tgg gag	gca gtt t	to acc g	gc ctc aca	cac ata	gac 5059
His Leu Glu F	he Trp Glu	Ala Val	he Thr G	ly Leu Thr	His Ile	Asp
1560		1565		1570		
get cat ttc o						
Ala His Phe I	eu Ser Gln	Thr Lys	Gln Ser (	ly Glu Asn	Phe Ala	Tyr
1575	158	0	158	5		
tta gta gcc t						
Leu Val Ala 7	Tyr Gln Ala	Thr Val	Cys Ala	Arg Ala Lys		Pro
1590	1595		1600		1605	
ceg tec tgg						
Pro Ser Trp /	Asp Val Met	Trp Lys	Cys Leu '	Thr Arg Leu	Lys Pro	Thr
	1610	16	15	162	•	
ctt gtg ggc						
Leu Val Gly	Pro Thr Pro	Leu Leu	Tyr Arg	Leu Gly Ser	Val Thr	Asn
162	5	1630		1635		
gag gtc acc						
Glu Val Thr	Leu Thr His	Pro Val	Thr Lys	Tyr Ile Ala	Thr Cys	Met
1640		1645		1650		
caa get gae						
Gln Ala Asp	Leu Glu Va	l Met Thr	Ser Thr	Trp Val Leu	Ala Gly	Gly
1655	16	60	16	65		
gtc tta gca	gcc gtc gc	e geg tat	tgc tta	geg acc ggg	tgt gtt	tec 5395

Va1	Leu	Ala	Ala	Val.	Ala	Ala	Tyr	Сув	Leu	Ala	Thr	Gly	Сув	Val	Ser	
167	0			167	5			1	680				168	5		
atc	att	ggc	cgt	tta	cac	atc	aac	çag	cga	gct	gtc	gtc	gct	ccg	gac	5443
Ile	11e	Gly	Arg	Leu	His	Ile	Asn	Gln	Arg	Ala	Va1	Val	Ala	Pro	Asp	
			169	0			1	695				170	0			
aag	gag	gtc	ctc	tat	gag	gct	ttt	gat	gag	atg	gag	gaa	tgt	gcc	tee	5491
ьув	Glu	Va1	Leu	Tyr	Glu	Λla	Phe	Asp	Glu	Met	Glu	Glu	Cys	Ala	Ser	
		17	05			:	1710				171	5				
aga	gcg	gct	ctc	ctt	gaa	gag	ggg	cag	cgg	ata	gcc	gag	atg	ctg	aag	5539
Arg	Ala	Ala	Leu	Leu	Glu	Glu	Gly	Gln	Arg	Ile	Ala	Glu	Met	Leu	Lys	
	1	720				1725				17	30					
						_	_		_			_	-	_	gac	5587
	-	Ile	Gln	Gly		Leu	Gln	Gln			Lys	Gln	Ala	Gln	Asp	
1	735				174	0			17	45						
			-	-		_	-								tgg	5635
		Pro	Ala			Ala	Ser	-		Lys	Met	Glu			Trp	
175	)			179	55			1	760				176	5		
																5683
-			_			Phe		-							gga	3003
AIA	гув	ніѕ	177	_	Asn	Pne		5er 775	GIA	ile	GIN	178		Ala	GIY	
			171				1	115				176				
ata	tan	202	ata		aaa	220	cot	aat	ata	aat	taa	ato	atro	aca	ttc	5731
_						Asn		-		-						5,01
220 tt	531	17			y		1790				179					
												-				
age	que	qee	ete	acc	agt	ccg	tta	tca	act	agc	acc	acc	atc	ctt	ctt	5779

Ser Ala Ala L	eu Thr Ser P	o Leu Ser Tl	ir Ser Thr Thr	Ile Leu Leu
1800	18	05	1810	
aac att ctg g	gg ggc tgg c	g geg tee e	a att gcg cca	ccc gcg ggg 5827
Asn Ile Leu G	ly Gly Trp L	eu Ala Ser G	n Ile Ala Pro	Pro Ala Gly
1815	1820		1825	
gcc act ggc t	tt gtt gtc a	gt ggc ctg g	g gga gct gct	gtt ggc agc 5875
Ala Thr Gly P	he Val Val S	er Gly Leu V	al Gly Ala Ala	Val Gly Ser
1830	1835	184	0	1845
ata ggc ttg g	gt aaa gtg c	tg gtg gac a	to etg gea ggg	tat ggt gcg 5923
Ile Gly Leu G	ly Lys Val L	eu Val Asp I	le Leu Ala Gly	Tyr Gly Ala
	1850	1855	186	50
ggc att tcg g	gg gcc ctc g	tc gcg ttt a	ag atc atg tc	ggc gag aag 5971
Gly Ile Ser G	ly Ala Leu V	al Ala Phe L	ys Ile Met Se	c Gly Glu Lys
1865	5	1870	1875	
ccc tcc atg g	ag gat gtc a	tc aac ttg c	tg cct ggg at	t ctg tct cca 6019
Pro Ser Met G	lu Asp Val I	le Asn Leu L	eu Pro Gly Ile	e Leu Ser Pro
1880	1	885	1890	
ggt gct ctg g	tg gtg gga g	tc atc tgc g	cg gcc att ct	g ege ege eat 6067
Gly Ala Leu V	al Val Gly V	al Ile Cys A	la Ala Ile Le	u Arg Arg His
1895	1900		1905	
gtg gga ccg g	gg gaa ggc g	eg gte caa t	gg atg aac ag	g ctt atc gcc 6115
Val Gly Pro G	ly Glu Gly A	1a Val Gln T	rp Met Asn Ar	g Leu Ile Ala
1910	1915	192	0	1925
ttc act tcc a	ıga gga aac c	ac gtc gcc c	ct act cac ta	c gtg acg gag 6163

Phe	Ala	Ser	Arg	Gly	Asn	нів	Val	Ala	Pro	Thr	His	Tyr	Val	Thr	G1u	
			193	0			1	935				1.940	)			
teg	gat	geg	tcg	cag	cgt	gtc	acc	caa	ctg	ctt	ggc	tot	ctc	act	ata	6211
ser	Asp	Ala	Ser	Gln	Arg	Va1	Thr	Gln	Leu	Leu	Gly	Ser	Leu	Thr	lle	
		19	4 5			1	950				1.95	5				
act	agt	cta	ctc	agg	aga	ctt	cac	aac	tgg	atc	act	gag	gat	tgc	ccc	6259
Thr	Ser	Leu	Leu	Arg	Arg	Leu	нів	Asn	Trp	Ile	Thr	Glu	Asp	Cys	Pro	
	15	960				1965				19	70					
atc	cca	tgc	gcc	ggc	tcg	tgg	ata	ege	gat	gtg	tgg	gac	tgg	gtc	tgt	6307
Ile	Pro	Сув	Ala	Gly	ser	Trp	Leu	Arg	Asp	Va1	Trp	Asp	Trp	Val	Cys	
1	975				198	0			19	85						
acc	atc	cta	aca	gac	ttt	aag	aac	tgg	ctg	acc	tcc	aag	ctg	ttc	cca	6355
Thr	Ile	Leu	Thr	Asp	Phe	Lys	Asn	Trp	Leu	Thr	ser	Lys	Leu	Phe	Pro	
199	0			199	5			2	000				200	5		
aag	atg	cct	ggc	ctc	aaa	ttt	atc	tct	tg¢	caa	aag	ggg	tac	aag	ggc	6403
Lys	Met	Pro	Gly	Leu	Pro	Phe	Ile	Ser	Cys	Gln	Lys	Gly	Tyr	гуз	Gλγ	
			20	10			2	015				202	0			
gtg	tgg	gcc	ggc	act	ggc	atc	atg	acc	aca	cga	tgc	ccc	tgc	ggc	gcc	6451
Va1	Trp	Ala	Gly	Thr	Gly	lle	Met	Thr	Thr	Arg	Cys	Pro	Сув	Gly	Ala	
		20	25				2030				20	3 5				
aac	atc	tct	ggc	aac	gtc	cgc	ttg	ggc	tet	atg	aga	atc	aca	gga	ccc	6499
Asn	Ile	Ser	Gly	Asn	Val	Arg	Leu	Gly	Ser	Met	Arg	Ile	Thr	Gly	Pro	
	2	040				2045	5			20	50					
aaa	acc	tgc	atg	aac	acc	tgg	cag	999	acc	ttt	cct	ato	aat	tgt	tat	6547

ьув	Thr	Сув	Met	Asn	Thr	тгр	Gln	Gly	Thr	Phe	Pro	11e	Asn	Сув	Tyr	
2	055				206	)			20	65						
aca	gaa	ggc	cag	tgc	ttg	ccg	aaa	ccc	gcg	tta	aac	ttc	aag	acc	gcc	6595
Thr	G1u	Gly	Gln	Cys	Leu	Pro	ьув	Pro	Ala	Leu	Asn	Phe	ьув	Thr	Ala	
2070				207	5			2	080				2085	õ		
atc	taa	aga	gtg	qcg	gac	tca	gag	tac	gcg	gaa	gtg	acg	cag	cac	gga	6643
		-						Tyr								
			209	0			2	095				210	)			
tas	tot	acc	tat	ata	aca	aaa	ct.a	acc	act	gac	aac	tta	aaa	qtc	cat	6691
								Thr								
	-2-	21					2110				211					
		2.1	05													
																6739
-															caa	6/39
Сув	Gln	Leu	Pro	ser	Pro			Phe	Ser			Asp	Gly	Val	GIn	
	2	120				2125	,			21	3 0					
atc	cat	agg	tcc	gcc	ccc	aca	cca	aag	ccg	ttt	ttc	cgg	gat	gag	gtc	6787
Ile	His	Arg	ser	Ala	Pro	Thr	Pro	Ъув	Pro	Phe	Phe	Arg	Asp	Glu	Val	
2	135				214	0			2	145						
tcg	ttc	agc	gtt	ggg	ctc	aat	tca	ttt	gtc	gtc	ggg	tct	cag	ctt	ccc	6835
Ser	Phe	ser	Va1	Gly	Leu	Asn	ser	Phe	Val	Val	Gly	ser	Gln	Leu	Pro	
2150	0			219	55			2	160				216	5		
tgt	gac	cct	gag	ccc	gac	act	gag	gta	gtg	atg	tcc	atg	cta	aca	gac	6883
Сув	Asp	Pro	Glu	Pro	Asp	Thr	Glu	val	Val	Met	ser	Met	Leu	Thr	Asp	
			21	70			:	175				218	0			
cca	tee	cat	at.c	acq	aca	gag	act	qca	qcq	cqq	cqt	tta	gcg	cgg	ggg	6931

Pro	ser	His	Ile	Thr	Ala	Glu	Ala	Ala	Ala	Arg	Arg	Leu	Ala	Arg	Gly	
		21	85			2	190				219	5				
tca	ccc	cca	tet	gag	gca	agc	tcc	tca	gcg	agc	cag	ctg	teg	gcg	cca	6979
ser	Pro	Pro	ser	Glu	Ala	Ser	ser	Ser	Ala	Ser	G1 n	Leu	ser	A1a	Pro	
	2:	00				2205				22	10					
teg	ctg	cga	gcc	acc	tgc	acc	acc	cac	ggt	agg	acc	tat	gat	gtg	gac	7027
Ser	Leu	Arg	Ala	Thr	Сув	Thr	Thr	His	Gly	Arg	Thr	Tyr	Asp	Val	Авр	
2	215				222	0			22	25						
atg	gtg	gat	gcc	aac	ctg	ttc	atg	aaa	ggc	ggc	gtg	att	cgg	ata	gag	7075
Met	Val	Asp	Ala	Asn	Leu	Phe	Met	Gly	Gly	Gly	Val	Ile	Arg	Ile	Glu	
2231	0			223	5			2	240				224	5		
tct	gag	tcc	aaa	gtg	gtc	gtt	ctg	gac	tcc	ctc	gac	tca	atg	acc	gag	7123
Ser	Glu	Ser	Lys	Val	Val	Val	Leu	Asp	ser	Leu	Asp	Ser	Met	Thr	Glu	
			22	50			2	255				226	0			
-															ccc	7171
Glu	Glu	Gly	Asp	Leu	Glu	Pro	Ser	Val	Pro	Ser	Glu	Tyr	Met	Leu	Pro	
		2 2	65				2270				227	75				
															tac	7219
Arg	Lys	Arg	Phe	Pro	Pro	Ala	Leu	Pro	Ala	Trp	Ala	Arg	Pro	Авр	Tyr	
	2	280				2285	ŝ			22	90					
aac	cca	ccg	ctt	gtg	gaa	teg	tgg	aag	agg	cca	gat	tac	caa	cca	ccc	7267
Asn	Pro	Pro	Leu	Val	Glu	Ser	Trp	Lys	Arg	Pro	Asp	Tyr	G1n	Pro	Pro	
2	2295				230	0			2	305						
act	gtt	gcg	ggc	tgt	get	ete	ccc	ccc	ccc	aaa	aag	acc	ccg	acg	cct	7315

Thr Val Ala Gly C	we als Lau Dro	Pro Pro Lvs Lvs	Thr Pro Thr Pro
	2315	2320	2325
2320			
cet eca agg aga e	ge egg aca gtg	ggt ctg agc gag	agc acc ata gga 7363
Pro Pro Arg Arg A			
2330	2	335	2340
gat gcc ctc caa c	ag ctg gcc atc	aag tcc ttt ggc	cag eec ccc cca 7411
Asp Ala Leu Gln G	ln Leu Ala Ile	Lys Ser Phe Gly	Gln Pro Pro Pro
2345	2350	235	5
age gge gat tea g			
Ser Gly Asp Ser G		Gly Ala Asp Ala 2370	Ala Asp Ser Gly
2360	2365	2370	
gat cgg aca ccc c	at asa asa tta	act ctt tea gag	aca ggt tct acc 7507
Asp Arg Thr Pro P			33
2375	2380	2385	•
tcc tcc atg ccc c	ec ctc gag ggg	gag cct ggg gac	cca gac ctg gag 7555
Ser Ser Met Pro P	ro Leu Glu Gly	Glu Pro Gly Asp	Pro Asp Leu Glu
2390	2395	2400	2405
cct gag cag gta g	gag ctt caa cct	cct ccc cag ggg	ggg gag gca gct 7603
Pro Glu Gln Val G	lu Leu Gln Pro	Pro Pro Gln Gly	-
2410	. 2	415	2420
			gag gag gat gac 7651
Pro Gly Ser Asp S	Ser Gly Ser Trp 2430		
2425	2430	24:	
tee gte gtg tac t	ge tee atg tea	tat tee tgg acc	ggg gct cta ata 7699
- 555			

Ser	Val	Val	Cys	Сув	Ser	Met	Ser	Tyr	Ser	Trp	Thr	Gly	Ala	Leu	Ile	
	2	140				2445				24	50					
act	cct	tgt	agc	ccc	gaa	gag	gaa	aag	ttg	cca	att	aac	tcc	ttg	age	7747
Thr	Pro	Cys	Ser	Pro	Glu	Glu	Glu	ьув	Leu	Pro	ıle	Asn	Ser	Leu	ser	
2	455				246	0			24	65						
aac	tcg	ctg	ttg	cga	tac	cat	aac	aag	gta	tac	tgt	act	aca	tca	aag	7795
Asn	ser	Leu	Leu	Arg	Tyr	His	Asn	ьув	Val	Tyr	Cys	Thr	Thr	ser	Lys	
247	0			247	5			2	480				248	5		
agt	gcc	tca	cta	agg	gct	aaa	aag	gta	act	ttt	gat	agg	atg	caa	gtg	7843
Ser	Ala	ser	Leu	Arg	Ala	Lys	ГАВ	Val	Thr	Phe	Asp	Arg	Met	Gln	Val	
			249	0			2	495				250	0			
cto	gac	gcc	tat	tat	gat	tca	gtc	tta	aag	gac	atc	aag	cta	gcg	gcc	7891
Leu	Asp	Ala	Tyr	Tyr	Asp	Ser	Va1	Leu	Lys	Asp	Ile	Lys	Leu	Ala	Ala	
		25	05			:	2510				251	5				
tcc	aag	gtc	agc	gca	agg	ctc	ctc	acc	tta	gag	gag	gcg	tgc	caa	ttg	7939
ser	ьув	Val	Ser	Ala	Arg	Leu	Leu	Thr	Leu	Glu	Glu	Ala	Cys	Gln	Leu	
	2	520				2525				25	3 0					
acc	cca	ccc	cac	tct	gca	aga	tcc	aag	tat	g gg	ttt	ggg	gct	aag	gag	7987
Thr	Pro	Pro	His	ser	Ala	Arg	ser	Lys	Tyr	Gly	Phe	Gly	Ala	Lys	Glu	
2	535				254	0			25	4 5						
gtc	cgc	agc	ttg	tcc	ggg	agg	gcc	gtc	aac	cac	atc	aag	tcc	gtg	tgg	8035
val	Arg	Ser	Leu	Ser	Gly	Arg	Ala	Val	Asn	His	Ile	Lys	ser	Val	Trp	
2550	0			255	5			2	560				256	5		
aag	gac	ctc	ttg	gaa	gac	tca	caa	aca	cca	att	cct	aca	acc	atc	atg	8083

Lys Asp Leu Le	u Glu Asp Ser	Gln Thr Pro	Ile Pro Thr	Thr Ile Met	
2	570	2575	258	0	
gcc aaa aat ga	g gtg ttc tgc	gtg gac ccc	gcc aag ggg	ggt aaa aaa	8131
Ala Lys Asn Gl	u Val Phe Cys	Val Asp Pro	Ala Lys Gly	сіу Був Був	
2585		2590	2595		
cca gct cgc ct	t atc gtt tac	cct gac ctc	ggc gtc agg	gtc tgc gag	8179
Pro Ala Arg Le	u Ile Val Tyr	Pro Asp Leu	Gly Val Arg	Val Cys Glu	
2600	260	5	2610		
aag atg gcc ct	t tat gat gto	aca caa aag	ctt cct cag	gcg gtg atg	8227
Lys Met Ala Le	u Tyr Asp Val	Thr Gln Lys	Leu Pro Glm	Ala Val Met	
2615	2620	2	625		
ggg gct tct ta	t ggc ttc cag	tac tee ecc	gct cag cgg	gtg gag ttt	8275
Gly Ala Ser Ty	r Gly Phe Gln	Tyr Ser Pro	Ala Gln Arg	Val Glu Phe	
2630	2635	2640		2645	
ctc ttg aag go	a tgg gcg gaa	aag aga gac	cct atg ggt	ttt tcg tat	8323
Leu Leu Lys Al	a Trp Ala Glu	Lys Arg Asp	Pro Met Gly	Phe Ser Tyr	
2	650	2655	266	0	
gat acc cga tg	c ttt gac tca	acc gtc act	gag aga gac	atc agg act	8371
Asp Thr Arg Cy	s Phe Asp Ser	Thr Val Thr	Glu Arg Asp	Ile Arg Thr	
2665		2670	2675		
gag gag tcc at	a tac cag gcc	tgc tcc tta	ccc gag gag	gee ega act	8419
Glu Glu Ser Il	e Tyr Gln Ala	Cys Ser Leu	Pro Glu Glu	Ala Arg Thr	
2680	268	5	2690		
gcc ata cac te	g ctg act gag	aga ctc tat	gtg gga ggg	ccc atg ttc	8467

Ala	Ile	His	Ser	Leu	Thr	Glu	Arg	Leu	Tyr	Val	Gly	Gly	Pro	Met	Phe	
2	695				270	)			27	05						
														agc Ser		8515
2710		гÀз	GIY	271		Cys	GIY		720	ura	Cyn	tur a	272		0.7	
gtg	ctt	acc	act	agt	atg	ggg	aac	acc	atc	aca	tgc	tat	gta	aaa	gcc	8563
Val	Leu	Thr	Thr	Ser	Met	Gly	Asn	Thr	Ile	Thr	Сув	Tyr	va1	Lys	Ala	
			273	0			2	735				274	0			
at a	aaa	act	tac	аап	act	aca	aaa	ata	att	aca	ccc	acq	atσ	ctg	gta	8611
														Leu		
		27					2750				275					
														gag		8659
Cys			Asp	Leu	Val	Val 2765		Ser	Glu	Ser 27		Gly	Thr	Glu	Glu	
	2	760				2/63	•			2,	70					
gac	gag	cgg	aac	ctg	aga	gcc	ttc	acg	gag	gct	atg	acc	agg	tat	tct	8707
Asp	Glu	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala	Met	Thr	Arg	Tyr	Ser	
2	775				278	0			2	785						
																8755
														cta Leu		6/55
279		FIO	GIY	27		110	ni g		800	~1-			280			
aca	tet	tgt	tcc	tca	aac	gtg	tet	gtg	gca	att	ggc	cca	cag	ggc	cgc	8803
Thr	Ser	Cys	Ser	Ser	Asn	Va1			Λla	Leu	Gly			Gly	Arg	
			28	1.0			2	815				282	0			
cgc	aga	tac	tac	ctg	acc	aga	gac	ccc	acc	act	tca	att	gcc	cgg	gct	8851

Arg Arg Tyr Tyr	Leu Thr Arg	Asp Pro Thr	Thr Ser Ile	Ala Arg Ala
2825		2830	2835	
gcc tgg gaa aca				
Ala Trp Glu Thr				Leu Gly Asn
2840	2845		2850	
atc atc cag tac				
Ile Ile Gln Tyr			Arg Met Val	Leu Met THI
2855	2860	26	.05	
cae tte tte too	ntt ata ata	262 262 222	acc cta gac	cag aac ctt 8995
His Phe Phe Ser				••• •• • • • • • • • • • • • • • • • •
2870	2875	2880	THE DOW HOP	2885
2070				
aac ttt gaa atg	tac qqa tcq	gtg tac tcc	gtg agt cct	ctg gac ctc 9043
Asn Phe Glu Met				
28		2895	2900	
cca gcc ata att	gaa agg tta	cac ggg ctt	gac gcc ttc	tet etg cac 9091
Pro Ala Ile Ile	Glu Arg Leu	His Gly Leu	Asp Ala Phe	Ser Leu His
2905		2910	2915	
aca tac act ccc	cac gaa ctg	acg cgg gtg	get tea gee	ctc aga aaa 9139
Thr Tyr Thr Pro	His Glu Leu	Thr Arg Val	Ala Ser Ala	Leu Arg Lys
2920	2925	5	2930	
ctt ggg gcg cca	ccc ctc aga	gcg tgg aag	agt egg geg	cgt gca gtt 9187
Leu Gly Ala Pro	Pro Leu Arg	Ala Trp Lys	Ser Arg Ala	Arg Ala Val
2935	2940	2	945	
agg g <b>cg</b> t <b>c</b> c ctc	atc tcc cgt	ggg ggg agg	geg gec gtt	tgc ggt cgg 9235

Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala Ala Val Cys Gly Arg	
2950 2955 2960 2965	
tad die tie dae tyg geg geg wag was ang oos ann	9283
Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu	
2970 2975 2980	
the transfer by the transfer of the transfer o	9331
ccg gag gca cgc ctc ctg gat ttg tcc agt tgg ttt acc gtc ggc gcc . S	,,,,,
2985 2990 2995	
2,703	
gge ggg ggc gac att tat cac age gtg teg egt gee ega ecc ege eta	9379
Gly Gly Asp Ile Tyr His Ser Val Ser Arg Ala Arg Pro Arg Leu	
3000 3005 3010	
tta ctc ctt agc cta ctc cta ctt tct gta ggg gta ggc ctc ttc cta	9427
Leu Leu Leu Ser Leu Leu Leu Ser Val Gly Val Gly Leu Phe Leu	
3015 3020 3025	
ctc ccc gct cga tag agcggcacac attagctaca ctccatagct aactgttcct	9482
Leu Pro Ala Arg	
3030	
	0542
ttttttttt tttttttt ttttttttt ttttttttt	3342
tottocotto toatottatt ctactttott tottggtggc tocatottag cootggtoac	9602
fortiogette featerfact eractitett rectigiogge reconcerna coorganismo	
ggotagotgt gaaaggtoog tgagoogoat gaotgoagag agtgoogtaa otggtototo	9662
gggggg	
tgcagatcat gt 9674	

```
<210> 6
<211> 3033
<212> PRT
<213> Hepatitis C virus
<400> 6
Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn
                              1.0
Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly
                           25
Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala
       3.5
Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro
                     55
                                      60
    50
Ile Pro Lys His Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly
                 70
                                   75
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp
                               90
              85
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro
          100
                          105
                                            110
Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys
                        120
                                         125
Gly Phe Ala Asp Leu Leu Gly Tyr Val Pro Val Val Gly Ala Pro Leu
                    135
                                      140
Ser Gly Val Ala Ser Ala Leu Ala His Gly Val Arg Val Leu Glu Asp
                 150
                                   155
                                                    160
Gly Val Asn Phe Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile
                              170
             165
Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Val
                                            190
                           185
Gln Val Lys Asn Thr Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser
                        200
                                         205
```

Asn	Asp	Ser	Ile	Thr	Trp	Gln	Leu	Glu	Ala	Ala	Va 1	Leu	His	Val	Pro
	210				219	,			2	20					
31y	Сув	Val	Pro	Cys	<b>G</b> lu	Lys	Met	Gly	Asn	Thr	Ser	Arg	Сув	Trp	Ile
225				23	0			2	235				240		
Pro	Val	Ser	Pro	Asn	Val	Ala	Val	Arg	Gln	Pro	Gly	Ala	Leu	Thr	Arg
			24	5			:	250				255			
Зlу	Leu	Arg	Thr	His	Ile	Asp	Met	Val	Val	Leu	Ser	Ala	Thr	Leu	Сув
		2	60				265				27	0			
Ser	Ala	Leu	Tyr	Val	Gly	Asp	Leu	Cys	Gly	Gly	Val	Met	Leu	Ala	Ser
	2	75				280				28	5				
31n	Met	Phe	Ile	Val	Ser	Pro	Gln	His	His	Trp	Phe	Val	Gln	Glu	Сув
	290				295	;			3	00					
Asn	Сув	ser	Ile	Tyr	Pro	Gly	Ala	Ile	Thr	Gly	His	Arg	Met	Ala	Trp
305				31	0			3	15				320		
Asp	Met	Met	Met	Asn	Trp	Ser	Pro	Thr	Thr	Thr	Met	Ile	Leu	Ala	Tyr
			32	5				330				335			
Val	Met	Arg	Val	Pro	Glu	Val	Ile	Ile	Asp	Ile	Ile	Ser	Gly	Ala	His
		3	40				345				35	0			
Frp	Gly	Val	Met	Phe	Gly	Leu	Ala	Tyr	Phe	Ser	Met	Gln	Gly	Ala	Trp
	3	55				360				36	5				
Ala	Lys	Val	Val	Val	Ile	Leu	Leu	Leu	Ala	Ser	Gly	Val	Asp	Ala	Tyr
	370				375	;			3	80					
rhr	Thr	Thr	Thr	Gly	Ser	Ala	Ala	Gly	Arg	Thr	Thr	Ser	Ser	Leu	Ala
385				3 9	0			3	95				400		
Ser	Ala	Phe	Ser	Pro	Gly	Ala	Arg	Gln	Asn	Ile	Gln	Leu	Ile	Asn	Thr
			40	5				410				415			
Asn	Gly	Ser	Trp	His	Ile	Asn	Arg	Thr	Ala	Leu	Asn	Сув	Asn	Asp	Ser
		4	20				425				43	0			
Leu	His	Thr	Gly	Phe	Phe	Thr	Ala	Leu	Phe	Tyr	Ile	His	ьуs	Phe	Asn
	4	3 5				440				44	5				
er	Ser	Gly	Сув	Pro	G1u	Arg	Leu	Ser	Ala	Сув	Arg	Asn	Ile	Glu	Asp
	450				455				4	60					

Phe	Arg	Ile	Gly	Trp	Gly	Ala	Leu	Gln	Tyr	Asp	Asp	Asn	Val	Thr	Asn
465				47	0			4	75				480		
Pro	Glu	Asp	Met	Arg	Pro	Tyr	Сув	Trp	His	Tyr	Pro	Pro	Lys	Gln	Cys
			48	5				190				495			
Gly	Va1	Val	Pro	Ala	Gly	Thr	Va1	Cys	Gly	Pro	Va1	Tyr	Сув	Phe	Thr
		5	00				505				51	)			
Pro	Ser	Pro	Val	Val	Va1	Gly	Thr	Thr	Asp	Arg	Leu	Gly	Val	Pro	Thr
	5	15				520				52	5				
Tyr	Thr	Trp	Gly	Glu	Asn	Glu	Thr	Asp	Val	Phe	Leu	Leu	Asn	ser	Thr
	530				535	5			5	40					
Arg	Pro	Pro	Ser	Gly	ser	Trp	Phe	Gly	Сув	Thr	Trp	Met	Asn	Ser	Thr
545				5 5	0			ţ	555				560		
Gly	Phe	Thr	Lys	Thr	Cys	Gly	Ala	Pro	Pro	Сув	Arg	Thr	Arg	Ala	Asp
			56	5				570				575			
Phe	Asn	Thr	Ser	Thr	Asp	Leu	Leu	Сув	Pro	Thr	qsA	Cys	Phe	Arg	Lys
		5	80				585				59	0			
His	Pro	Glu	Ala	Thr	Tyr	Ile	Lys	Сув	Gly	Ser	Gly	Pro	Trp	Leu	Thr
	5	95				600				60	15				
Pro	ьys	Сув	Leu	Val	Asp	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Сув
	610				615	5			6	20					
Thr	Val	Asn	Tyr	Ser	Thr	Phe	Lys	Ile	Arg	Met	Tyr	Val	Gly	Gly	Val
625				63	0			6	35				640		
Glu	His	Arg	Leu	Met	Ala	Ala	Сув	Asn	Phe	Thr	Arg	Gly	Asp	Arg	Сла
			64	5				650				655			
Asn	Leu	Glu	Asp	Arg	Авр	Arg	Ser	Gln	Gln	Thr	Pro	Leu	Leu	His	Ser
		6	60				665				67	0			
Thr	Thr	Glu	Trp	Ala	Ile	Leu	Pro	Cys	ser	Phe	Ser	Авр	Leu	Pro	Ala
	6	75				680				68	5				
Leu	Ser	Thr	Gly	Leu	Leu	His	Leu	His	Gln	Asn	Ile	Va 1	Asp	Va1	Gln
	690				699	5			7	00					
туr	Met	Tyr	Gly	Leu	Ser	Pro	Ala	Leu	Thr	Gln	туг	Ile	Val	Arg	Trp
705				71	.0			1	715				720		

Glu	Trp	Val	Va1	Leu	Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Val	Сув
			72	5				130				735			
Ala	Cys	Leu	Trp	Met	Leu	Ile	Leu	Leu	Gly	Gln	Ala	Glu	Ala	Ala	Leu
		7.	4 0				745				75	0			
Glu	Lys	Leu	Val	Va1	Leu	His	Ala	Ala	Ser	Ala	Ala	ser	Cys	Asn	Gly
	7	55				760				76	5				
Phe	Leu	Tyr	Phe	Va1	Ile	Phe	Leu	Val	Ala	Ala	Trp	His	Ίle	ьув	Gly
	770				775				7	80					
Arg	Va1	Val	Pro	Leu	Ala	A1a	Tyr	ser	Leu	Thr	Gly	Leu	Trp	Pro	Phe
785				79	0			7	95				800		
Сув	Leu	Leu	Leu	Leu	Ala	Leu	Pro	Gln	Gln	Ala	Tyr	Ala	Tyr	Asp	A1a
			8 0	5				310				815			
Ser	va 1	нів	Gly	Gln	Val	Gly	Ala	Ala	Leu	Leu	Val	Leu	Ile	Thr	Leu
		8	20				825				83	0			
Phe	Thr	Leu	Thr	Pro	Gly	Tyr	Lys	Thr	Leu	Leu	Ser	Gln	Ser	Leu	Trp
	8	135				840				8 4	5				
Trp	Leu	Сув	Tyr	Leu	Leu	Thr	Leu	Ala	Glu	Thr	Met	Va1	Gln	Glu	Trp
	850				855	i			8	60					
Ala	Pro	Ser	Met	Gln	Ala	Arg	Gly	Gly	Arg	Asp	Gly	Ile	Ile	Trp	Ala
865				87	0			4	375				880		
Ala	Thr	Ile	Phe	Cys	Pro	Gly	Val	Va1	Phe	Asp	Ile	Thr	Lys	Trp	Leu
			88	5				890				895			
Leu	Ala	Val	Leu	Gly	Pro	Gly	Tyr	Leu	Leu	Arg	Gly	Ala	Leu	Thr	Arg
		9	00				905				91	0			
Val	Pro	Tyr	Phe	Va1	Arg	Ala	нів	Ala	Leu	Leu	Arg	Met	Сув	Thr	Met
	9	15				920				92	25				
Val	Arg	нів	Leu	Ala	Gly	Gly	Arg	Tyr	Val	Gln	Met	Ala	Leu	Leu	Ala
	930				935	i			9	40					
Leu	Gly	Arg	Trp	Thr	Gly	Thr	Tyr	11e	Tyr	Asp	His	Leu	Thr	Pro	Met
945				95	0				955				960		
ser	Asp	Trp	Ala	Ala	ser	Gly	Leu	Arg	Asp	Leu	Ala	Val	Ala	Val	Glu
			96	5				970				975			

Pro Ile Ile	Phe Ser	Pro Met	Glu Lys	Lys Val	Ile Val	Trp Gly Ala
9	80		985		990	
Glu Thr Ala	Ala Cys	Gly Asp	Ile Leu	His Gly	Leu Pro	Val Ser Ala
995		1000		10	05	
Arg Leu Gly	Λrg Glu	Ile Leu	Leu Gly	Pro Ala	Asp Gly	Tyr Thr Ser
1010		1015		1020		
Lys Gly Trp	Lys Leu	Leu Ala	Pro Ile	Thr Ala	Tyr Ala	Gln Gln Thr
1025	10	3 0		L035		1040
Arg Gly Leu	Leu Gly	Ser Ile	Val Val	Ser Met	Thr Gly	Arg Asp Lys
	1045		1050		105	5
Thr Glu Gln	Ala Gly	Glu Val	Gln Val	Leu Ser	Thr Val	Thr Gln Ser
10	60		1065		1070	
Phe Leu Gly	Thr Ser	Ile Ser	Gly Val	Leu Trp	Thr Val	Tyr His Gly
1075		1080	)	10	85	
Ala Gly Asn	Lys Thr	Leu Ala	Gly Ser	Arg Gly	Pro Val	Thr Gln Met
1090		1095		1100		
Tyr Ser Ser	Ala Glu	Gly Asp	Leu Val	Gly Trp	Pro Ser	Pro Pro Gly
1105	11:	10		1115		1120
Thr Lys Ser	Leu Glu	Pro Cys	Thr Cys	Gly Ala	Val Asp	Leu Tyr Leu
	1125		1130		113	5
Val Thr Arg	Asn Ala	Asp Val	Ile Pro	Ala Arg	Arg Arg	Gly Asp Lys
11	4 0		1145		1150	
Arg Gly Ala	Leu Leu	Ser Pro	Arg Pro	Leu Ser	Thr Leu	Lys Gly Ser
1155		1160	)	11	.65	
Ser Gly Gly	Pro Val	Leu Cys	Pro Arg	Gly His	Ala Val	Gly Ile Phe
1170		1175		1180		
Arg Ala Ala	Val Cys	Ser Arg	Gly Val	Ala Lys	Ser Ile	Asp Phe Ile
1185	11:	90		L195		1200
Pro Val Glu	Thr Leu	Asp Ile	Val Thr	Arg Ser	Pro Thr	Phe Ser Asp
	1205		1210		121	5
Asn Ser Thr	Pro Pro	Ala Val	Pro Glr	Thr Tyr	Gln Val	Gly Tyr Leu
12	20		1225		1230	

uio	Δla	Pro	Thr	Glv	Ser	Glv	LVS	Ser	Thr	Lvs	Val	Pro	Va1	Ala	Tyr
1123		135		017		1240				12					•
Ala			Glv	Tvr	Lvs	Val	Leu	Val	Leu	Asn	Pro	Ser	Val	Ala	Ala
	250			•	125					60					
		Gly	Phe	Gly	Ala	Tyr	Leu	ser	Lys	Ala	His	Gly	Ile	Asn	Pro
1269	5			127	0			1	275				1280	0	
Asn	Ile	Arg	Thr	Gly	va1	Arg	Thr	Val	Thr	Thr	Gly	Glu	Pro	Ile	Thr
			128	5			1	290				1299	5		
Tyr	Ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala	Авр	Gly	Gly	Cys	Ala	Gly	Gly
		13	00			:	1305				131	0			
Ala	Tyr	Asp	Ile	Ile	Tle	Cys	Asp	Glu	Cys	нів	Ser	Val	Asp	Ala	Thr
	13	15				1320				13	25				
Thr	Ile	Leu	Gly	Ile	Gly	Thr	Va1	Leu	Asp	Gln	Ala	Glu	Thr	Ala	Gly
1	330				133	5			13	40					
Val	Arg	Leu	Thr	Val	Leu	Ala	Thr	Ala	Thr	Pro	Pro	Gly	ser	Val	Thr
134	õ			135	0			1	355				136	ð	
Thr	Pro	His	Pro	Asn	Ile	Glu	Glu	Val	Ala	Leu	Gly	Gln	Glu	Gly	Glu
			136	5			1	370				137	5		
Ile	Pro	Phe	Tyr	Gly	Arg	Ala	Phe	Pro	Leu	Ser	Tyr	Ile	Lys	Gly	Gly
		13	80			:	1385				139	0			
Arg	His	Leu	Ile	Phe	Сув	His	Ser	Lys	Lys	Lys	Cys	Asp	Glu	Leu	Ala
	1	95				1400	)			14	05				
Thr	Ala	Leu	Arg	Gly	Met	Gly	Leu	Asn	Ala	Val	Ala	Tyr	Tyr	Arg	Gly
1	410				141	5			14	120					
Leu	Asp	Val	Ser	Ile	Ile	Pro	Thr	Gln	Gly	Asp	Val	Val	Val	Val	Ala
142	ö			143	80			1	435				144	0	
Thr	Asp	Ala	Leu	Met	Thr	Gly	туг	Thr	Gly	Asp	Phe	Asp	Ser	Va1	Ile
			144	5			1	450				145	5		
Asp	Cys	Asn	Val	Λla	Val	Thr	Gln	Ala	Val	Asp	Phe	Ser	Leu	Asp	Pro
		14	60			:	1465				147	0			
Thr	Phe	Thr	Ile	Thr	Thr	Gln	Thr	Val	Pro	Gln	Asp	Ala	Val.	Ser	Arg
	Phe Thr Ile Thr Thr					1480	)			14	85				

Ser	Gln	Arg	Arg	Gly	Arg	Thr	G1y	Arg	Gly	Arg	Leu	Gly	Ile	Tyr	Arg
1	490				149	5			15	0 0					
Tyr	Val	Ser	Thr	Gly	Glu	Arg	Ala	Ser	Gly	Met	Phe	Asp	Ser	Val	Val
1509	5			151	.0			1	515				1520	)	
Leu	Сув	Glu	Сув	Tyr	Asp	Ala	G1y	Ala	Ala	ттр	Tyr	Glu	Leu	Ser	Pro
			152	5			1	530				153	5		
Val	Glu	Thr	Thr	Val	Arg	Leu	Arg	A1a	Tyr	Phe	Asn	Thr	Pro	Gly	Leu
		15	4 0			1	545				155	0			
Pro	Va1	Cys	Gln	Asp	His	Leu	Glu	Phe	Trp	Glu	Ala	۷al	Phe	Thr	Gly
	15	555				1560				15	65				
Leu	Thr	нів	Ile	Asp	Ala	His	Phe	Leu	Ser	Gln	Thr	Lys	Gln	Ser	Gly
1	570				157	5			15	80					
Glu	Asn	Phe	Ala	Tyr	Leu	Val	Ala	Tyr	Gln	Ala	Thr	Val	Сув	Ala	Arg
1585	j.		1590 la Pro Pro Pro S					1	595				1600	)	
Ala	Lys	Ala	Pro	Pro	Pro	Ser	Trp	Asp	Val	Met	Trp	Lys	Сув	Leu	Thr
			160	5			1	610				161	5		
Arg	Leu	Lys	Pro	Thr	Leu	Val	Gly	Pro	Thr	Pro	Leu	Leu	Tyr	Arg	Leu
		16	20			1	625				163	0			
Gly	Ser	Val	Thr	Asn	Glu	Va1	Thr	Leu	Thr	His	Pro	Val	Thr	ГÀЗ	Tyr
	16	535				1640				16	4 5				
Ile	Ala	Thr	Cys	Met	Gln	Ala	Asp	Leu	Glu	Val	Met	Thr	Ser	Thr	Trp
1	650				165	5			16	60					
Val	Leu	Ala	Gly	Gly	Val	Leu	Ala	Ala	Val	Ala	Ala	Tyr	Cys	Leu	Ala
1665	5			167	0			1	675				1680	)	
Thr	Gly	Сув	Val	ser	Ile	Ile	Gly	Arg	Leu	His	Ile	Asn	Gln	Arg	Ala
			168	5			1	690				169	5		
Val	Val	Ala	Pro	Asp	Lys	Glu	Val	Leu	Tyr	Glu	Ala	Phe	Asp	Glu	Met
		1700			1	1705				171	0				
Glu	Glu	Сув	A1a	ser	Arg	Ala	Ala	Leu	Leu	Glu	Glu	Gly	Gln	Arg	Ile
	13	715				1720				17	25				
A1a	G1u	Met	Leu	Lys	ser	Lys	Ile	Gln	G1y	Leu	Lец	Gln	G1n	Ala	Ser
1	730	31u Met Leu Lys 30			173	5			17	40					

Lys	Gln	A1a	$_{\tt Gln}$	Asp	Ile	G1n	Pro	Ala	Val	Gln	Ala	Ser	Trp	Pro	Lys
1745				175	0			1	755				1760	)	
Met	G1 u	Gln	Phe	Trp	Ala	ьуs	ніѕ	Met	Trp	Asn	Phe	I1e	Ser	Gly	Ile
			176	5			1	770				1775	ő		
Gln	туг	Leu	Ala	Gly	Leu	Ser	Thr	Leu	Pro	Gly	Asn	Pro	A1a	Val	Ala
		17	80			1	785				179	0			
ser	Met	Met	Ala	Phe	ser	Ala	Ala	Leu	Thr	Ser	Pro	Leu	Ser	Thr	Ser
	17	95				1800				180	5				
Thr	Thr	Ile	Leu	Leu	Asn	Ile	Leu	Gly	Gly	Trp	Leu	Ala	ser	Gln	Ile
1	810				181	5			18	20					
Ala	Pro	Pro	Ala	Gly	Ala	Thr	Gly	Phe	Val	Val	ser	Gly	Leu	Val	Gly
1825	;			183	0			1	835				184	)	
Ala	Ala	val	Gly	Ser	Ile	Gly	Leu	Gly	ьув	Val	Leu	Val	Asp	Ile	Leu
			184	5			1	850				185	5		
Ala	Gly	Tyr	Gly	Ala	Gly	Ile	ser	Gly	Ala	Leu	Val	Ala	Phe	Lys	Ile
		18	60			;	1865				187	0			
Met	Ser	Gly	Glu	Гуз	Pro	Ser	меt	Ģlu	Asp	Val	Ile	Asn	Leu	Leu	Pro
	1	375				1880				18	8 5				
Gly	Ile	Leu	ser	Pro	Gly	Ala	Leu	Val	Val	Gly	Val	Ile	Сув	Ala	Ala
1	890				189	5			19	00					
Ile	Leu	Arg	Arg	нів	Val	Gly	Pro	Gly	Glu	Gly	Ala	Val	Gln	Trp	Met
1905	5			191	L 0			1	915				192	0	
Asn	Arg	Leu	Ile	Ala	Phe	Ala	ser	Arg	Gly	Asn	His	Val	Ala	Pro	Thr
			192	25			1	930				193	5		
His	Tyr	Va 1	Thr	Glu	ser	Авр	Ala	Ser	Gln	Arg	Va 1	Thr	Gln	Leu	Leu
		19	40				1945				195	0			
Gly	ser	Leu	Thr	Ile	Thr	ser	Leu	Leu	Arg	Arg	Leu	нів	Asn	Trp	Ile
	1	955				1960				19	65				
Thr	Glu	Asp	Cys	Pro	Ile	Pro	Сув	Ala	Gly	Ser	Trp	Leu	Arg	Asp	Val
1	970				197	5			1	980					
Trp	Asp	Trp	Va1	Cys	Thr	Ile	Leu	Thr	Asp	Phe	Lys	Asn	Trp	Leu	Thr
1989	5			199	90			1	995				200	0	

Ser	Lуs	Leu	Phe	Pro	ьув	Меt	Pro	Gly	Leu	Pro	Phe	Ile	Ser	Cys	Gln
			200	5			2	010				2015	5		
Lys	Gly	Tyr	Lys	Gly	va1	Trp	Ala	Gly	Thr	Gly	ıle	Met	Thr	Thr	Arg
		20	20			2	025				203	0			
Cys	Pro	Cys	Gly	Ala	Asn	Ile	Ser	Gly	Asn	Va1	Arg	Leu	Gly	Ser	Met
	20	35				2040				20	4.5				
Arg	Ile	Thr	Gly	Pro	Lys	Thr	Сув	Met	Asn	Thr	Trp	Gln	Gly	Thr	Phe
2	050				205	5			20	60					
Pro	Ile	Asn	Cys	туr	Thr	Glu	Gly	Gln	Сув	Leu	Pro	Lys	Pro	Ala	Leu
2065	;			207	0			2	075				2080	)	
Asn	Phe	Lys	Thr	Ala	Ile	Trp	Arg	Val	Ala	Ala	Ser	Glu	Tyr	Ala	Glu
			208	15			2	090				2099	5		
Val	Thr	Gln	His	Gly	ser	Tyr	Ala	Tyr	Ile	Thr	Gly	Leu	Thr	Thr	qaA
		21	00			:	105				211	0			
Asn	Leu	Lys	Va1	Pro	Сув	Gln	Leu	Pro	Ser	Pro	Glu	Phe	Phe	Ser	Trp
	2	115				2120				21	25				
val	Asp	Gly	Val	Gln	Ile	нів	Arg	Ser	Ala	Pro	Thr	Pro	Lys	Pro	Phe
2	130				213	5			21	40					
Phe	Arg	Asp	Glu	Val	Ser	Phe	ser	Val	Gly	Leu	Asn	Ser	Phe	Val	Va1
2145	5			215	50			2	155				216	0	
Gly	Ser	Gln	Leu	Pro	Cys	Asp	Pro	Glu	Pro	Asp	Thr	Glu	Val	Val	Met
			216	55			2	170				217	5		
Ser	Met	Leu	Thr	Asp	Pro	ser	His	Ile	Thr	Ala	G1 u	Ala	Ala	Ala	Arg
		21	80			:	185				219	0			
Arg	Leu	Ala	Arg	Gly	Ser	Pro	Pro	ser	Glu	Ala	Ser	Ser	Ser	Ala	ser
	2	195				2200				22	05				
Gln	Leu	Ser	Ala	Pro	Ser	Leu	Arg	Ala	Thr	Сув	Thr	Thr	His	Gly	Arg
2	210		2215			5			22	220					
Thr	Tyr	Asp	Val.	Asp	Met	Val	Asp	Ala	Asn	Leu	Phe	Met	Gly	Gly	Gly
2225	2230				3 0			2	235				224	0	
Val	Ile	Arg	ıle	Glu	ser	Glu	Ser	Lys	Val	Val	Va1	Leu	Asp	Ser	Leu
			229	5			2	250				225	5		

Asp	Ser	Met	Thr	Glu	Glu	Glu	Gly	Asp	Leu	Glu	Pro	Ser	Val	Pro	Ser
		22	60			2	265				227	0			
Glu	Tyr	Met	Leu	Pro	Arg	Lys	Arg	Phe	Pro	Pro	Ala	Leu	Pro	Ala	Trp
	22	275				2280				22	B 5				
Ala	Arg	Pro	Asp	Tyr	Asn	Pro	Pro	Leu	Val	Glu	ser	Trp	Lys	Arg	Pro
2	290				229	5			23	00					
Asp	Tyr	Gln	Pro	Pro	Thr	va1	Ala	Gly	Сув	Ala	Leu	Pro	Pro	Pro	гла
2309	5			231	0			2	315				2320	)	
Lys	Thr	Pro	Thr	Pro	Pro	Pro	Arg	Arg	Arg	Arg	Thr	va1	Gly	Leu	Ser
			232	5			2	330				2335	5		
Glu	Ser	Thr	Ile	Gly	Asp	Ala	Leu	Gln	Gln	Leu	Ala	Ile	ГÅа	ser	Phe
		23	40			2	345				235	0			
Gly	Gln	Pro	Pro	Pro	Ser	Gly	Asp	Ser	Gly	Leu	Ser	Thr	Gly	Ala	qaA
	23	355				2360				23	65				
Ala	Ala	Asp	Ser	Gly	Asp	Arg	Thr	Pro	Pro	Asp	Glu	Leu	Ala	Leu	Ser
2	370				237	5			23	80					
Glu	Thr	Gly	ser	Thr	Ser	Ser	Met	Pro	Pro	Leu	Glu	Gly	Glu	Pro	Gly
2389	5			239	0			2	395				2400	)	
Asp	Pro	Asp	Leu	Glu	Pro	Glu	Gln	Val	Glu	Leu	Gln	Pro	Pro	Pro	Gln
			240	5			2	410				2415	5		
Gly	Gly	Glu	Ala	Ala	Pro	Gly	Ser	Asp	Ser	Gly	Ser	Trp	Ser	Thr	Сув
		24	20			2	425				243	0			
ser	Glu	Glu	Asp	Asp	Ser	Val	Val	Сув	Сув	Ser	Met	ser	Tyr	Ser	Trp
	24	135				2440				24	45				
Thr	Gly	Ala	Leu	Ile	Thr	Pro	Сув	ser	Pro	Glu	Glu	Glu	ьуз	Leu	Pro
2	450				245	5			24	60					
Ile	Asn	Ser	Leu	ser	Asn	Ser	Leu	Leu	Arg	Tyr	His	Asn	Lys	Val	Tyr
2465	5	2470			0			2	475				2480	)	
Cys	Thr	Thr	Ser	Lys	Ser	Ala	Ser	Leu	Arg	Ala	Lys	гув	Val	Thr	Phe
			248	5			2	490				249	ô		
Asp	Arg	Met	Gln	Va1	Leu	Asp	Ala	Tyr	Tyr	Asp	ser	Val	Leu	гув	Asp
		rg Met Gln Val				2	505				251	0			

Ile Lys Leu	Ala Ala Sc	r Lys Val	Ser Ala	Arg Leu I	Leu Thr Leu Glu
2515		2520		2525	
Glu Ala Cys	Gln Leu Th	r Pro Pro	His Ser	Ala Arg S	Ser Lys Tyr Gly
2530	25	35	25	40	
Phe Gly Ala	Lys Glu Va	l Arg Ser	Leu Ser	Gly Arg A	Ala Val Asn His
2545	2550		2555		2560
Ile Lys Ser	Val Trp Ly	s Asp Leu	Leu Glu	Asp Ser (	In Thr Pro Ile
	2565	2	570	2	2575
Pro Thr Thr	Ile Met Al	a Lys Asn	Glu Val	Phe Cys V	/al Asp Pro Ala
25	80	2585		2590	
Lys Gly Gly	гла гла ы	o Ala Arg	Leu Ile	Val Tyr I	Pro Asp Leu Gly
2595		2600		2605	
Val Arg Val	Cys Glu Ly	s Met Ala	Leu Tyr	Asp Val 1	Thr Gln Lys Leu
2610	20	15	26	20	
Pro Gln Ala	Val Met Gl	y Ala Ser	Tyr Gly	Phe Gln 7	Tyr Ser Pro Ala
2625	2630		2635		2640
Gln Arg Val	Glu Phe Le	u Leu Lys	Ala Trp	Ala Glu I	Lys Arg Asp Pro
	2645		650		2655
Met Gly Phe	Ser Tyr As	p Thr Arg	Cys Phe	Asp Ser 1	Thr Val Thr Glu
26	60	2665		2670	
Arg Asp Ile	Arg Thr Gl	u Glu Ser	Ile Tyr	Gln Ala (	Cys Ser Leu Pro
2675		2680		2685	
Glu Glu Ala	Arg Thr Al	a Ile His	Ser Leu	Thr Glu A	Arg Leu Tyr Val
2690	26	95	27	00	
Gly Gly Pro	Met Phe As	n Ser Lys	Gly Gln	Ser Cys (	Bly Tyr Arg Arg
2705	2710		2715		2720
Cys Arg Ala	Ser Gly Va	l Leu Thr	Thr Ser	Met Gly A	Asn Thr Ile Thr
	2725	2	730		2735
Cys Tyr Val	Lys Ala Le	u Ala Ala	Сув Гув	Ala Ala G	Hy Ile Ile Ala
27	40	2745		2750	
Pro Thr Met	Leu Val Cy	s Gly Asp	Asp Leu	Val Val J	le Ser Glu Ser
2755		2760		2765	

Gln	Gly	Thr	Glu	Glu	Asp	Glu	Arg	Asn	Leu	Arg	Ala	Phe	Thr	Glu	Ala
2	770				277	5			27	80					
Met	Thr	Arg	Tyr	Ser	Ala	Pro	Pro	Gly	Asp	Pro	Pro	Arg	Pro	Glu	Tyr
2785	i			279	0			2	795				2800	ı	
Asp	Leu	Glu	Leu	lle	Thr	Ser	Сув	Ser	ser	Asn	Va 1	Ser	Val	Ala	Leu
			280	5			2	810				281	5		
gly	Pro	Gln	Gly	Arg	Arg	Arg	Tyr	Tyr	Leu	Thr	Arg	Asp	Pro	Thr	Thr
		28	20			2	825				283	0			
ser	11e	Ala	Arg	Ala	Ala	Trp	Glu	Thr	Val	Arg	His	ser	Pro	Val	Asn
	28	35				2840				28	45				
Ser	Trp	Leu	Gly	Asn	Ile	Ile	Gln	Tyr	Ala	Pro	Thr	Ile	Trp	Val	Arg
2	850				285	5			28	60					
Met	Val	Leu	Met	Thr	His	Phe	Phe	Ser	Ile	Leu	Met	Ala	Gln	Asp	Thr
2865	5			287	0			2	875				288	ນ	
Leu	Asp	Gln	Asn	Leu	Asn	Phe	Glu	Met	Tyr	Gly	ser	Va 1	Tyr	Ser	Val
			288	35			2	890				289	5		
ser	Pro	Leu	Asp	Leu	Pro	Ala	Ile	Ile	Glu	Arg	Leu	ніз	Gly	Leu	Asp
		29	00			:	2905				291	0			
Ala	Phe	Ser	Leu	His	Thr	Tyr	Thr	Pro	His	Glu	Leu	Thr	Arg	Val	Ala
	25	15				2920				29	25				
Ser	Ala	Leu	Arg	Lys	Leu	Gly	Ala	Pro	Pro	Leu	Arg	Ala	Trp	Lys	Ser
2	930				293	5			29	40					
Arg	Ala	Arg	Ala	Val	Arg	Ala	Ser	Leu	Ile	Ser	Arg	Gly	Gly	Arg	Ala
2945	5			295	50			2	955				296	0	
Ala	Val	Сув	Gly	Arg	Tyr	Leu	Phe	Asn	Trp	Ala	Val	Lys	Thr	ьуs	Leu
			296	55			2	970				297	5		
Lys	Leu	Thr	Pro	Leu	Pro	Glu	Ala	Arg	Leu	Leu	Asp	Leu	Ser	Ser	Trp
		29	80			:	2985				299	0			
Phe	Thr	Val	Gly	Ala	Gly	Gly	Gly	Авр	Tle	Tyr	нія	Ser	Val	Ser	Arg
	25	995				3000	1	,		3 0	05				
Ala	Arg	Pro	Arg	Leu	Leu	Leu	Leu	ser	Leu	Leu	Leu	Leu	Ser	Val	Gly
3	010				301	5			3 (	20					

```
3025
                3030
<210> 7
<211> 8024
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: replicon
<400> 7
accugecceu aauaggggeg acacueegee augaaucaeu eeccugugag gaacuacugu 60
cuucacgcag aaagegecua gecauggegu uaguaugagu gueguacage cuccaggeec 120
coccuccog ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360
ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420
egeagguucu eeggeegeuu ggguggagag geuauuegge uaugaeuggg cacaacagae 480
aaucggcugc ucugaugceg ceguguuceg geugucageg caggggegee egguucuuuu 540
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600
guggeuggee acgaegggeg uuccuugege ageugugeue gaeguuguea eugaageggg 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccugcegag aaaguaucca ucauggeuga ugcaaugegg eggeugeaua egeuugauce 780
ggouaccugo ccauucgaco accaagogaa acaucgcauc gagogagoac guacucggau 840
ggaagceggu cuuguegaue aggaugaucu ggaegaagag caucagggge uegegecage 900
egaacuguuc gecaggeuca aggegegeau gecegaegge gaggaucueg uegugaecea 960
uggegaugee ugeuugeega auaucauggu ggaaaaugge egeuuuucug gauucaucga 1020
```

Val Gly Leu Phe Leu Leu Pro Ala Arg

cuguggeegg cugggugugg eggacegcua ucaggacaua geguuggeua ecegugauau 1080

ugcugaagag cuuggeggeg aaugggeuga cogcuuceuc gugcuuuacg guaucgeecgc 1140 uccegauucg cagegeaucg ceuucuaucg ceuucuugac gaguucuucu gaguuuaaac 1200 ccucucccuc coccccccu aacgunacug gcogaagcog cuuggaauaa ggcoggugug 1260 cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320 aaccuggecc ugucuucuug acgagcauuc cuaggggucu uuccccucuc gccaaaggaa 1380 ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440 caacgucugu agcgacccuu ugcaggcagc ggaacccccc accuggcgac aggugccucu 1500 geggecaaaa gecaegugua uaagauacae cugcaaagge ggeacaacee cagugecaeg 1560 uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620 ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680 caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggccccccg aaccacgggg 1740 acgugguuun ccuuugaaaa acacgangan accanggene ccancacuge unangeecag 1800 caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860 caggoogggg aaguccaaau couguccaca gucucucagu couuccucgg aacaaccauc 1920 ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980 gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gcccagcccc 2040 ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100 cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160 ccgagaccca uuucgaccuu gaaggggucc ucgggggggc cggugcucug cccuaggggc 2220 cacquequug ggeueuuceg ageageugug ugeucueggg geguggeeaa auceaucgau 2280 uucauccccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340 acgccaccgg cugugcccca gaccuaucag gucggguacu ugcaugcucc aacuggcagu 2400 ggaaagagca ccaagguccc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460 aaccccucgg uagcugccac ccugggguuu ggggcguacc uauccaaggc acauggcauc 2520 aaucccaaca uuaggacugg agucaggace gugaugaceg gggaggccau cacguacucc 2580 acauauggca aauuucucgc cgauggggc ugcgcuagcg gcgccuauga caucaucaua 2640 ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700 caagcagaga cagccggggu cagacuaacu gugcuggcua cggccacacc ccccggguca 2760 gugacaaccc cccaucccga uauagaaaag guaggccucg ggcgggaggg ugagaucccc 2820 uucuauggga gggcgauucc ccuauccugc aucaagggag ggagacaccu gauuuucugc 2880 cacucaaaga aaaaguguga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940 guggcauacu auagaggguu ggacgucucc auaauaccag cucagggaga uguggugguc 3000

guogecaceg aegeceucau gaegggguae aeuggagaeu uugaeucegu gaucgaeuge 3060 aauguagegg ucacccaage uguegacuuc agocuggacc ccaccuucac uauaaccaca 3120 cagacuguec cacaagacge ugucucaege agucagegee gegggegeac agguagagga 3180 agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240 quaquqcuuu gugaguqcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300 accaceguca ggcuuagage guauuucaac aegeeeggee uaeeegugug ucaagaeeau 3360 cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucucc 3420 caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480 qccaqaqcca aggccccucc cccquccugg gacgccaugu ggaagugccu ggcccgacuc 3540 aagccuacgc nugcgggccc cacaccucuc cuguaccguu ugggcccuau uaccaaugag 3600 gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660 gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720 cuggegacug gaugeguuuc caucaucgge egeuugeacg ucaaccageg aguegueguu 3780 gegeeggaua aggaggueeu guaugaggeu uuugaugaga uggaggaaug egeeueuagg 3840 geggeueuca ucgaagaggg geageggaua geegagaugu ugaagueeaa gauceaagge 3900 uugcugcagc aggccucuaa gcaggcccag gacauacaac ccgcuaugca ggcuucaugg 3960 cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020 cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080 geogeccuea ecaguecquu guegaecagu accaecauce uucucaacau caugggagge 4140 uggunagegu cccagaucgc accaccegeg ggggccaccg gcunngucgu caguggccug 4200 gugggggcug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260 uauggugegg geauuueggg ggeeeuegue geauueaaga ueaugueugg egagaageee 4320 ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380 qqqqucaucu gcqcqqccau ucuqcqccqc cacquqqqac cqqqqqaqqq cqcqquccaa 4440 uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500 quqacqqaqu cgqaugcquc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560 agccuacuca gaagacucca caauuggaua acugaggacu gccccauccc augcuccgga 4620 uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680 cugaccucua aauuguuccc caagcugecc ggccuccccu ucaucucuug ucaaaagggg 4740 uacaagggug ugugggeegg cacuggcauc augaccaege geugeeeuug eggegeeaac 4800 aucucuggea augucegeeu gggeueuaug aggaucacag ggeeuaaaac eugeaugaac 4860 accuggoagg ggaccuuucc uaucaauugc uacacggagg gccagugogc gccgaaaccc 4920

cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980 cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aauuccuugc 5040 caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100 cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160 gcuqucgggu cccagcuucc cugugaaccu gagcccgacg cagacguauu gagguccaug 5220 cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280 ccuccaucug aggcgagcuc cucagugagc cagcuaucag caccguegcu gegggccacc 5340 ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400 ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460 auggeegagg aagagagega ceuugageee ucaauaceau eggagugeau geueeecagg 5520 ageggguuue caegggeeuu aceggeuugg geaeggeeug acuacaacee geegeuegug 5580 gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucccccc 5640 cccaagaagg ccccgacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700 accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760 ggugaugcag gcucguccac gggggcgggc gccgccgaau ccggcggucc gacguccccu 5820 ggugagccgg cccccucaga gacagguucc gccuccucua ugcccccccu cgaggggag 5880 ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucccc ccaggggggg 5940 gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000 accgugugcu geuccauguc auacuccugg accggggeuc uaauaacucc cuguagcccc 6060 gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120 guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180 acgcaagugc ucgacgccca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240 aaggucageg caaggeuccu caccuuggag gaggegugee aguugacuce accecauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 genegecuca negunnacce ngacenegge gueeggguen gegagaaaan ggeeencuan 6540 gacauuacac aaaagcuucc ucaggeggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggecug cucccugece gaggaggece geacugecau acacuegeug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uageggccug caaggcugeg gggauaguug egcccacaau gcugguaugc 6960 ggcaaugacc uaguagucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020 agagecunca eggaggecau gaccagguac ucugececue euggugauce ecceagaceg 7080 qaauauqacc uqqaqcuaau aacauccuqu uccucaaaug ugucuguggc guugggcccg 7140 eggggeegee geagauacua ceugaeeaga gaceeaacea cuccaeuege eegggeugee 7200 ugggaaacag uuagacacuc cccuaucaau ucauggcugg gaaacaucau ccaguaugcu 7260 ccaaccauau ggguucgcau gguccuaaug acacacuucu ucuccauucu caugguccaa 7320 qacacccuqg accagaaccu caacuuugag auguauggau caguauacuc cgugaauccu 7380 uuqqaccuuc cagccauaau ugagagguua cacqggcuug acgccuuuuc uaugcacaca 7440 uacucucace acquacugae geqgguggeu ucageccuca gaaaacuugg ggegecacec 7500 cucagggugu ggaagagucg ggcucgcgca gucagggcgu cccucaucuc ccguggaggg 7560 aaageggeeg uuugeggeeg auaucucuuc aauugggegg ugaagaccaa geucaaacuc 7620 acuccauuge eggaggegeg ceuacuggac uuauccaguu gguucacegu eggegeegge 7680 gggggcgaca uuuuucacag cgugucgegc gcccgacccc gcucauuacu cuucggccua 7740 cuccuacuuu ucquaggggu aggccucuuc cuacuccccg cucgguagag cggcacacac 7800 uuuuuuuu cuuuuuuuu uuuuuccuc uuucuuccu ucucaucuua uucuacuuuc 7920 uuucuuqquq gcuccaucuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980 8024 augacuqcaq aqaquqccqu aacuqgucuc ucugcagauc augu

```
<210> 8
<211> 7994
<212> RNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: replicon

<400> 8
accugeccu aauaggggeg acacucegce augaaucacu ecceugugag gaacuacugu 60

cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120 cccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180 aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 egeuugegag ugeeeeggga ggueueguag acogugeace augageacaa auccuaaace 360 ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420 egeagguucu eeggeegeuu ggguggagag geuauuegge uaugaeuggg cacaacagae 480 aaucggcuge ucugaugceg ceguguuceg geugucageg cagggggggce egguucuuuu 540 ugucaagacc gaccuguceg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600 guggeuggee acgaegggeg uuccuugege ageugugeue gaeguuguea eugaageggg 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugeegag aaaguaucea ucauggeuga ugeaaugegg eggeugeaua egeuugauce 780 ggcuaccuge ccauucgace accaagegaa acaucgcaue gagegageae guacucggau 840 ggaageeggu cuuguegaue aggaugaueu ggaegaagag caucagggge uegegeeage 900 ogaacuguuc gocaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggcgaugec ugcuugeega auaucauggu ggaaaaugge egcuuuucug gauucaucga 1020 cuguggeegg cugggugugg eggacegeua ucaggacaua geguuggeua ecegugauau 1080 ugcugaagag cuuggeggeg aaugggeuga cegeuuceue gugcuuuaeg guaucgeege 1140 ucccgauucg cagegeaucg ceuucuaucg ceuucuugac gaguucuucu gaguuuaaac 1200 ccucucccuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260 cquuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320 aaccuggccc ugucuucuug acgagcauuc cuaggggucu uuccccucuc gccaaaggaa 1380 ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440 caacgucugu agcgacccuu ugcaggcagc ggaacccccc accuggcgac aggugccucu 1500 geggecaaaa gecaegugua uaagauacae eugeaaagge ggeacaacee cagugecaeg 1560 uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620 ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680 caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggccccccg aaccacgggg 1740 acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800 caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860 caggeegggg aaguecaaau ceuguecaca gueucucagu ecuuceucgg aacaaccauc 1920 ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980 gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gcccagcccc 2040 ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100 cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160 ccgagaccca uuucgaccuu gaaggggucc ucgggggggc cggugcucug cccuaggggc 2220 cacgucguug ggcucuuccg agcagcugug ugcucucggg gcguggccaa auccaucgau 2280 uucauccccg uugagacacu cgacguuguu acaaggucuc ccacuuucag ugacaacagc 2340 acgccaccgg cugugccca gaccuaucag gucggguacu ugcaugcucc aacuggcagu 2400 ggaaagagca ccaagguccc ugucgcguau gccgcccagg gguacaaagu acuagugcuu 2460 aaccccucgg uagcugccac ccugggguuu ggggcguacc uauccaaggc acauggcauc 2520 aaucccaaca uuaggacugg agucaggacc gugaugaccg gggaggccau cacguacucc 2580 acauauggca aauuucucgc cgaugggggc ugcgcuagcg gcgccuauga caucaucaua 2640 ugcgaugaau gccacgcugu ggaugcuacc uccauucucg gcaucggaac gguccuugau 2700 caagcagaga cagcoggggu cagacuaacu gugcuggcua cggccacacc ccccggguca 2760 gugacaaccc cccaucccga uauagaagag guaggccucg ggcgggaggg ugagaucccc 2820 uucuauggga gggcgauucc ccuauccugc aucaagggag ggagacaccu gauuuucugc 2880 cacucaaaga aaaaguguga cgagcucgcg gcggcccuuc ggggcauggg cuugaaugcc 2940 guggcauacu auagaggguu ggacgucucc auaauaccag cucagggaga uguggugguc 3000 guegecaceg aegeceucau gaegggguae aeuggagaeu uugaeucegu gauegaeuge 3060 aauguagegg ucacccaage uguegacuuc ageeuggace ecaccuucae uauaaccaca 3120 cagacuguce cacaagacge ugucucaege agucagegee gegggegeae agguagagga 3180 agacagggca cuuauaggua uguuuccacu ggugaacgag ccucaggaau guuugacagu 3240 guagugcuuu gugagugcua cgacgcaggg gcugcguggu acgaucucac accagcggag 3300 accaccguca ggcuuagagc guauuucaac acgcccggcc uacccgugug ucaagaccau 3360 cuugaauuuu gggaggcagu uuucaccggc cucacacaca uagacgccca cuuccucuc 3420 caaacaaagc aagcggggga gaacuucgcg uaccuaguag ccuaccaagc uacggugugc 3480 gecagageca aggeeceuce cocquecugg gacgecaugu ggaagugecu ggeecgacue 3540 aagecuaege uugegggeee cacaccucuc cuguaeeguu ugggeecuau uaccaaugag 3600 gucacccuca cacacccugg gacgaaguac aucgccacau gcaugcaagc ugaccuugag 3660 gucaugacca gcacgugggu ccuagcugga ggaguccugg cagccgucgc cgcauauugc 3720 cuggegacug gaugeguuue caucauegge egeuugeaeg ucaaccageg aguegueguu 3780 gcgccggaua aggagguccu guaugaggcu uuugaugaga uggaggaaug cgccucuagg 3840 geggeueuca uegaagaggg geageggaua geegagaugu ugaagueeaa gaueeaagge 3900

uugcugcage aggecucuaa geaggeeeag gacauacaac eegcuaugca ggeuucaugg 3960 cccaaagugg aacaauuuug ggccagacac auguggaacu ucauuagcgg cauccaauac 4020 cucgcaggau ugucaacacu gccagggaac cccgcggugg cuuccaugau ggcauucagu 4080 geegeecuca ecagueeguu guegaecagu accaccauce uucucaacau caugggagge 4140 ugguuagegu cccagaucge accaccegeg ggggccaccg gcuuugucgu caguggccug 4200 gugggggcug ccgugggcag cauaggccug gguaaggugc ugguggacau ccuggcagga 4260 uauggugcgg gcauuucggg ggcccucguc gcauucaaga ucaugucugg cgagaagecc 4320 ucuauggaag augucaucaa ucuacugccu gggauccugu cuccgggagc ccugguggug 4380 ggggucaucu gcgcggccau ucugcgccgc cacgugggac cgggggaggg cgcgguccaa 4440 uggaugaaca ggcuuauugc cuuugcuucc agaggaaacc acgucgcccc uacucacuac 4500 gugacggagu cggaugcguc gcagcgugug acccaacuac uuggcucucu uacuauaacc 4560 agecuacuea gaagacueca caauuggaua acugaggacu geeccauece augeucegga 4620 uccuggcucc gcgacgugug ggacuggguu ugcaccaucu ugacagacuu caaaaauugg 4680 cugaccucua aauuguuccc caagcugccc ggccuccccu ucaucucuug ucaaaagggg 4740 uacaagggug ugugggccgg cacuggcauc augaccacgc gcugcccuug cggcgccaac 4800 aucucuggca augucegecu gggcucuaug aggaucacag ggccuaaaac cugcaugaac 4860 accuggragg ggaccunucc uaucaanngc nacacggagg gccagugcgc gccgaaaccc 4920 cccacgaacu acaagaccgc caucuggagg guggcggccu cggaguacgc ggaggugacg 4980 cagcaugggu cguacuccua uguaacagga cugaccacug acaaucugaa aauuccuugc 5040 caacuaccuu cuccagaguu uuucuccugg guggacggug ugcagaucca uagguuugca 5100 cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuau 5160 geuguegggu eccageuuce eugugaaceu gageeegaeg cagaeguauu gaggueeaug 5220 cuaacagauc cgccccacau cacggcggag acugcggcgc ggcgcuuggc acggggauca 5280 ccuccaucug aggogagouc cucagugago cagcuaucag caceguegou gegggecace 5340 ugcaccaccc acagcaacac cuaugacgug gacauggucg augccaaccu gcucauggag 5400 ggcggugugg cucagacaga gccugagucc agggugcccg uucuggacuu ucucgagcca 5460 auggoogagg aagagagoga couugagooc ucaauaccau oggagugoau gouccocagg 5520 ageggguuuc caegggecuu aeeggeuugg geaeggeeug acuacaacee geegeuegug 5580 gaaucgugga ggaggccaga uuaccaaccg cccaccguug cugguugugc ucucccccc 5640 cccaagaagg ccccgacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700 accauaucag aageccucca geaacuggee aucaagaccu uuggecagee ecceuegage 5760 ggugaugcag gcucquccac gggggggggc gccgccgaau ccggcggucc gacguccccu 5820

ggugagcegg cececucaga gacagguuce gecuccucua ugecececcu egaggggag 5880 ccuggagauc cggaccugga gucugaucag guagagcuuc aaccuccccc ccaggggggg 5940 gggguageue cegguueggg eueggggueu uggueuaeuu geueegagga ggaegauaee 6000 accgugugcu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagcccc 6060 qaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120 ququacuqua caacaucaaa qagcgccuca cagagggcua aaaagguaac uuuugacagg 6180 acgcaaguge ucgacgccca uuaugacuca gucuuaaagg acaucaagcu ageggeuucc 6240 aaggucageg caaggcuccu caccuuggag gaggegugec aguugacucc accccauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 genegecuca negunnacce ngacenegge guceggguen gegagaaaan ggeeencuan 6540 gacauuacac aaaagcuucc ucaggoggua augggagcuu ccuauggcuu ccaguacucc 6600 ccuqcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucquaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggcoug cuccougoco gaggaggeoc gcacugcoau acacucgoug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugec gegecagegg ggugcuaacc acuageaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggaggacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugecceuc cuggugaucc ecccagaceg gaauaugace uggageuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140 gacccaacca cuccacuege cegggeugee ugggaaacag uuagacacue cecuaucaau 7200 ucauggcugg gaaacaucau ccaguaugcu ccaaccauau ggguucgcau gguccuaaug 7260 acacacuucu ucuccauucu cauqquccaa gacacccugg accagaaccu caacuuugag 7320 auquauqqau caquauacuc cquqaauccu uuggaccuuc cagccauaau ugagagguua 7380 cacgggcuug acgccuuuuc uaugcacaca uacucucacc acgaacugac geggguggcu 7440 ucageceuca gaaaacuugg ggegeeacce eucagggugu ggaagagueg ggeuegegea 7500 gucagggcgu cccucaucuc ccguggaggg aaagcggccg uuugcggccg auaucucuuc 7560 aauugggegg ugaagaccaa geucaaacuc acuccauuge eggaggegeg ceuacuggac 7620 uuauccaguu gguucaccgu cggcgccggc gggggcgaca uuuuucacag cgugucgcgc 7680 geeegaceee geucauuacu cuucggeeua cuccuacuuu ucguaggggu aggeeucuuc 7740

```
cuacuccocg cucgguagag cggcacacac uagguacacu ccauagcuaa cuguuccuuu 7800
ишишишиш ишишишиш ишишишиши ишишишиши сишишиши ишишисские 7860
uuucuucccu ucucaucuua uucuacuuuc uuucuuggug gcuccaucuu agcccuaguc 7920
acggcuagcu gugaaagguc cgugagccgc augacugcag agagugccgu aacuggucuc 7980
                                                          7994
ucugcagauc augu
<210> 9
<211> 340
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
<400> 9
accugacca aauagggga acaaucaga augaaucacu accaugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccuccc qgagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
                                                              340
egeuugegag ugeceeggga ggueueguag accgugeace
<210> 10
<211> 340
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
```

```
<400> 10
accegecceu aauaggggeg acacucegee augaaucacu ceccugugag gaacuacugu 60
cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugeeeg gecauuuggg egugeeeeeg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
ugcuugcgag ugccccggga ggucucguag accgugcacc
                                                     340
<210> 11
<211> 236
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
<400> 11
uauucuacuu ucuuucuuqq uggcuccauc uuagcccuag ucacggcuag cugugaaagg 180
uccgugagec geaugacuge agagagugec guaacuggue ucucugeaga ucaugu 236
<210> 12
<211> 232
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
```

```
<400> 12
ageggeacae auuageuaca euceauageu aacuguuceu uuuuuuuuuu uuuuuuuuu 60
uuuuuuuuu uuuuuuuu uuuuuuuu uuuccucuu ucuucccuu ucaucuuauu 120
cuacuuucuu ucuugguggc uccaucuuag cccuggucac ggcuagcugu gaaagguccg 180
ugageegeau gaeugeagag agugeeguaa euggueueue ugeagaucau gu
<210> 13
<211> 17
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 13
                                                          17
cgggagagcc atagtgg
<210> 14
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 14
                                                          19
agtaccacaa ggcctttcg
<210> 15
```

```
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 15
                                                            21
ctgcggaacc ggtgagtaca c
<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 16
                                                            20
aacaagatgg attgcacgca
<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 17
cgtcaagaag gcgatagaag
                                                            20
```

```
<210> 18
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 18
gcactototg cagtcatgeg getcacggac
                                                             30
<210> 19
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 19
cccctgtgag gaactactgt cttcacgc
                                                            28
<210> 20
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
```

```
<223> Description of Artificial Sequence: synthetic DNA
<400> 20
                                                            24
ccgggagagc catagtggtc tgcg
<210> 21
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 21
                                                             30
ccactcaaag aaaaagtgtg acgagctcgc
<210> 22
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 22
                                                           18
ggettgggca cggectga
<21.0> 23
<211> 30
```

```
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 23
                                                              30
goggtgaaga ccaagetcaa actcactcca
<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 24
agaacctgcg tgcaatccat c
                                                             21
<210> 25
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 25
                                                             23
cccgtcatga gggcgtcggt ggc
```

```
<210> 26
<211> 27
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 26
accagcaacg gtgggcggtt ggtaatc
                                                             27
<210> 27
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 27
ggcacgcgac acgctgtg
                                                            18
<210> 28
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
```

```
<223> Description of Artificial Sequence: synthetic DNA
<400> 28
agctagccgt gactagggct aagatggagc
                                                             30
<210> 29
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:synthetic
     DNA(primer)
<400> 29
aacaagatgg attgcacgca
                                                            20
<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:synthetic
    DNA(primer)
<400> 30
                                                            20
cgtcaagaag gcgatagaag
```

```
<210> 31
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:synthetic DNA
<400> 31
geactetetq caqteatqcq getcacqqac
                                                             30
<210> 32
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:synthetic DNA
<400> 32
                                                             28
eccetgtgag gaactactgt etteacge
<210> 33
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
```

```
<400> 33
                                                            24
eegggagage catagtggte tgeg
<210> 34
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 34
                                                             30
ccactcaaag aaaaagtgtg acgagetege
<210> 35
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic
     DNA (primer)
<400> 35
ggcttgggca cggcctga
                                                           18
<210> 36
<211> 30
<212> DNA
```

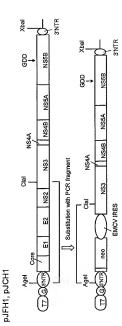
```
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 36
                                                             30
geggtgaaga ceaageteaa acteacteca
<210> 37
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 37
                                                            21
agaacctgcg tgcaatccat c
<210> 38
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 38
cccgtcatga gggcgtcggt ggc
                                                            23
```

```
<210> 39
<211> 27
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 39
                                                             27
accagcaacg gtgggcggtt ggtaatc
<210> 40
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 40
                                                           18
ggaacgcgac acgctgtg
<210> 41
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
```

<400> 41 agctagccgt gactagggct aagatggagc

30

[Title of Document] Drawings [Figure 1]



pSGREP-JFH1, pSGREP-JCH1

#### [Figure 2A]

ADDUGGGGGG	ANUNGGGGGG	ACACUOOGO	AUGAAUCAC	50 CCCCUGUGAG	GAACUACUSID
TO CUUCACGCAG				J10 GUCGUNCAGO	
130	140 GGAGAGCCAU	AGUGGUCUGO	GGAACÇGGUG	AGDACACOSG	180 AAUUXCO969
190 Ambacugggu	200 CCUUUCUU93	AUAAACCCAC	220 UCUNUGCCC	GCCVAGAGG 330	00000000000000000000000000000000000000
250 CAAGACUGCU	260 AGCOGAGUAG	COURSOCKING	088 COSAAAGOCU	290 UGUGGUACUG	ODEKUARKEX
310 0000000330	320 UGCCCCGGGA	330	340 ACCGUGGACC	350 AUGAGCAÇAA	066 Corakuzoua
370 UCANAGANAN	380 ACCAAAACAA	390 39000000000000000000000000000000000	UCGCCCAAUG	AUUGAACAAG	420 ADOUUAGOUA
430 OSCASGUUCU	COSSCOSCUU	450 Geguegagaga	GCUNUUCGGC	UAUGACUGGG	480 Cacaacagac
AADOGGGGGG	500 UCUGAUGCOS	510 CCCUGUUCCG	520 GCUGUCAGOS	CAGGGGGGGG	540 CGGÜUCUURU
950 UGUCAAGACC	9ACCUCUCOG	57G GUGCCCUGAA	. 580 UGAACUGCAG	SACGAGGCAG	600 OGCGGCUAUC
610 GUGGCUGGCC .	ACGACGGGGG	UNGCONGCCC	NGCUGUGCUC NGCUGUGCUC	GACGUUGUCA 4	660 CUCANGCOGG
AAGGGACUGG	680 CUGCIJAUUGG	698 GCGAAGUGCC	700 GGGGCAGGAU	CUCCUGUCAU I	720 CUCACCUUGC
			OCCANOCOGG	770 CGGCGCAUA	
GCUACCUGC	008 00A00UGAGG	VCGVÝCOGYV 610	ACAUCGCAUC	GAGCGAGCAC C	840 WACUCGGAD
GGARGCÓGGU	860 CUUGUCGAUC	870 AGGAUGAUCU	980 GGACGAAGAG	CAUCAGGGGC L	900 DOGOGCCAGC
910 CGAACUGUUC	920 BCCAGGCUCA	930 10000000000	GCCCGACGGC	950 GAGGAUCUCG U	
UGGCGAUGCC 1	980 UGCUNGCCEA	990 BUAUCAUGGU	1900 GGAAAAUGGC	COCCUUUCUG G	AUUCAUCGA
COGUGGCCGG C	1040 CUGGGUGUGG	1050 CGGACCGCUA	1060 UCAGGACAUA	GOGUUGGCUA C	1089 COGIXIAUAU
1090 USCUGARGEG	1100 CUUGGOGGGG	L110 AAUGGGCUGA	COSCUUCCUC	1130 GUGCUUUACG G	1140 UNUOGCOGC
UCCCGAUUCG C	L160	1170 20000UAU03		BYGONCOCCO 6 7136	
1319	1220	. 1230 NACOUUNCUS	1240 GCCGAAGCCG	1250 DUUGGBAUAA G	1260 GCCGGUGUG
1270 CGUJUGUCUA U	1280 JAUGUUUUUUU	1290 KCACCAUAU	1300 DGCCGUCUOU 1	1310 IGGCAAUGUG A	1320 SGGCCCGGA
1330 AACCUGGCCC L	1340 KIUCUKICUUG J	1350 CCAGCAUK:	1360 CUAGGGGUCU T	UCCCCUCUC G	1380 Changgaa

# [Figure 2B]

1390 IXICAAGGUCU	1400 GUNGAAUGUC	1410 GUGAAGGAAG	1420 CAGUUCCUCU	GGAAGCUUCU	144 UGAAGACAA
1450 CARCIGUCUGU	1460 AGCGACCCUU	UGCAGGCAGC	1480 GGAACCCCCC	1490 ACCUGGOGÁC	150 AGGUGCCUCI
1510 GOGGCCAAAA	1520 GCCACGUGUA	1530 UANGAUACAC	1640 CUSCAAAGGC	1550 GGCACARCCC	1560 CAGUGOCAC
1570 DOGRAGING	1580 GAUAGUUGUG	1590 GANAGAGUCA	1600 ANUGGCUCUC	CUCAAGOGUA	UDCANCAAGE
1610 GGCUGAAGGA	1640 UGCCCAGAAG	1650 GUACCCCAUU	1660 Guaugggauc	1670 UGMÜCUGGGG	1690 20000000000
1690 CAUGCUUUAC	AUGUGUUUAG	1710 AKUUDDKGCU	. 1720 AAAAACGUCU	1730 AGGCCCCCCG	AACCACGGG
1750 Acguscuuu	1760 CCUUKGAAAA	1770 ACACGAUGAU	ACCAUGGCIC	CCAUCACUGC	UUAUGCCCAG
	eccuccuese	CCCCNUAGUG	B 40 ADUAUDA DUDA DUD		
	arguccaaru	CCOGUCCACA	GUCUCICAGU	ccnnccncea	AACAACCAUC
			1960 GCUGGCAACA		
			GAGGGGGACU	nacovaccna 5030	
		GCCGUGCAAG	UGUGGNGCCG	LICGACCUAUA	OCOGGOCACG
OGGAACGCUG 2170	AUGUCAUCCC 2180		2140 CGCGGGACA 2200	AGCGGGGGGGC 2210	AUDOCUCUCO 2720
COGRGACICA 2230	DUDOGACCUD 2240	GAAGGGGGCC 2250	UCGGGGGGGC	COGUSCUCUS 2270	CCCVAGGGGG
CACGUCGUUG 2290	GGCUCUUCCG 2300	AGCAGCUGUG 2310	UGCUCUCGGG 2320	GCGUGGCCAA 2330	AUCCAUCGAU 2340
DOCAUCOCOG	UUGAGACACU	CGACGUUGUU	ACAAGGUCUC	CCYCACACAC	UGACAACAGG
ACCCCACCGG	CUGUGCCCCA 2420	GACCUAUCAG 2430	2380 GUCGGGUACU 2440	VGCAUGCUCC 2450	AACUGGGAGU 2460
OGAAAGAGCA 2470	CCARGGUCCC 2480	UGUOGCGUAU 2490	GCCGCCCAGG	GGUACAĀAGU 2510	acuaguscaiu
AACCCCUCGG 2530	UAGCUGCCAC 2540	2550	GGGGCGUACC	UAUCCAÃGGC 2570	acauggcauc 2580
ANUCCCARCA	UUAGGACUGG	AGUCAGGACC	GUGAUGACCG	GGGAGGCCAU	CACGUACUCC
2590 ACAUAUGGCA			2628 UGOGCUNGOG		
2650 UGCGAUGAAU	GCCACGCUGU	2670 GGAUGCUACC	nccannence \$880	2690 GCAUCGGAAC	CCCCCCCOON 2700
2710 CAAGCAGAGA	2720 CAGCCGGGGU	2730 CAGACUAACU	2740 GUGCUGGCUA	2750 COGCCACACC	2760 CCCCCGGGUCA

#### [Figure 2C]

GUONCANCO	2780 CCCAUCCCGA	2790 Unuagaagag	2800 GUAGGCCUOG	2810 GGCGGGAGGG	282 UGAGAUCCC
2830 UUCUAUGGG/	2840 GGGCGATUCC	2850 CCUNUÇEUĞÜ	AUCAAGGAG AUCAAGGAG	2870 GGAGACACCU	
CACUCARAGA	2900 Aaaaguguga	2910 CGAGCÚCSCS	2920 GCGGCCCUDG	2930 GGGGCAUGGG	CUUGARUGO
S950 CHOCCADACU	2960 NUNGAGGGGUU	2970 GGACGUCUCC	2980 AUAAUACCAG	2990 CUCAGGGAGA	300 UQUQQUQQU
3010 GDCGCCACC3	ACROCCUCAU	GACGGGGGAC	3040 ACUGGAGACU	3050 UUGACUCCGU	3050 GAUCGACUG
3070 AADGUAGOOG	DCACOCAAGC	UGUCGACUUC	AGCCEGRACC	CCACCOUCAC	
CAGACOGUCO	CACAAGACGC	USUCUCACGO	AGUCAGOGOC	GOGGGGGGAC	
	CUUNUNGGUA	rennnecyen	GGUGAACGAG		
	GUGAGAXICUA	COACGCAGGG	gchdocheen	ACGAUCUCAC	
	GGCUUAGAGC	GUAUUUCAAC	ACCCCCCCCCC		
		3390 000CACCGGC 3450	3400 CUCACACACA 3460	3420 UAGACGCCA 3470	3420 CUUCCUCUCU 3480
3430 CAAACAAAGC 3490	AAGOGGGGGA 3500	GAACUUCGCG 3510	UACCUAGUAG 3520	CCUACCAAGC:	DACGGUGUGC 3540
GOCAGAGOCA 3550	AGGCCCCUCC	ccccuccues 3570	GACGCCÄÜĞÜ 3580	GGAAGUGCCU ( 3590	26000000000000000000000000000000000000
ANGCOUNCEC	1000000000 1620	3630	CUGBACCEDD 3640	UGGGCCCUAU !	DACCAADGAG 3660
3670	CACACCCUGG 3660	GACGAAGUAC 3690	AUGGCCACAU 3700	GCAUGCAAGC 1 3710	GACCUUGAG 3720
3730	3740	CCUAGCUGGA 3750	3760	3770	3780
3790	3800	CAUCAUCGGC 3810	3820	. 3830	3840
3850	3860	GUAUGAGGCU 3870	3880	3890	3900
3910	3920	GCAGCGGAUA 3930	3940	3950	3950
3970	3980	GCAGGCCCAG 4 3990	4000	4010	4020
CCCAAAGUGG 4030	AACAAUUUUG (	GGCCAGACAC A	AUGUGGAACU 1 4060	CAUUAGOGG C	
CUCGCAGGAU	UGUCAACACU	GCCAGGGAAC (	cccccccooo	MUCCAUGAU G	GCAUUCAGU
4090 GCCGCCCCCA	4100 CCAGUCOGUU (	4110 GKCACCAĞU	ACCACCAUCC L	HKUCAACAU C	4140 AUGOGAGGC

#### [Figure 2D]

4150 UGGUUAGÇGU	4160 CCCAGAUCGC	4170 ACCACCOGGG	4180 GGGGCCXCCG	4190 GCUDUGUCGU	CAGUGGCCU
4210 GUGGGGGCUG	4220 COGUGGGCAG	4230 CAUAGGCCUG	4240 GBUANGGUGU	4250 UGGUGGACAU	ccuggexgoi
4270 UAUGGUGCGG	4289 GCAUUUCGGG	4290 GGCCCUCGUC	4300 GCAUUCAAGA		
43.30 UCUAUGGAAG	4340 AUGUCAUCAA	4350 UCUACUGOCU	4360 GGGAIXCCUGU	4370 CUCCGGGAGC	4380 CCUGGUGGU
4390 GGGDCAUCU	4400 GOGOGGCCAU	4410 UCUGOGOGG	CACGUGGGAC	4430 CGGGGGAGGG	GCGGGUCCAF
6450 USGAUGAACA	4460 GGCUUAUUGC	4470 CUUUGCUUCC	4480 Agaggaaacc	ACGUCGCCCC	UACUCACUAC
4510 GUXTACGGAGU	4520 CGGAUGCGUC	4530 GCAGCGUGUG	4540 ACCCAACUAC	4550 UNGGCUCUCU	UACUAUAACO
AGCCUNCUCA	4560 GAGACUXCA	4590 Caauuggaua	ACUCIAGGACU ACUCIAGGACU	GCCCCAUCCC	AUGCUCCGGA
4630 UCCUGGCUCC	4640 GCGACGUGUG	4650 GGACUGGGUU	UGCACCAUCU	4670 DGACAGACUU	CANADAUUGG
	ANUUUUUUCCC	CYNGCUGCCC	escencecen	ECAUCUCUUG	UCANANGGGG
4750 UACAAGGGUG		CACUGGCAUC			
		GGGCUCUAUG	AGGAUCACAG	GCCCUANAC	CUGCAUGNAC
		4090 DDODKADUKU			
	,	CAUCUGGAGG			
		5010 UGUAACAGGA	CUGACCACUG		
		5070 UUUCUCCKIG			
		CCGGGAUGAG			
		5198 CUGUGAACCO			
		E250 CACGGCGGAG	•		
5290 COUCCAUCUG	AGGCGAGCUC	5310 CUCAGUGAGC	S330 CAGCUAUCAG	CACCGUCGCU	5340 GOGGGCCACC
5350 UGCACCACCC	5360 ACAGCAACAC	5370 CUAUGACGUG	OBEZ DÓUDOUADAD	5390 AUGCCAACCU	5400 BCUCAUGOAG
5410 GGCGGUGUGG	5420 CUCAGACAGA	5430 GCCUGAGUCC	5440 AGGGUGCCOG 1	5450 UUCUGGACUU	5460 UCUCGAGCCA
5470 AUGGCCGAGG	5480 AAGAGAGCGA	5490 CCUUGAGÇÇÇ	5500 UKSSAUACKSU	5510 CGGAGUGCAU	5520 SCUCCCCAGG

#### [Figure 2E]

5530 AGCCCG0000	5540 CACGGGCCUU	5550 ACCGGCUUGG	SS60 GCACGGCCUG	5570 ACUACAROCO	5580 GCCGCUCGUCI
5590 GANDOGDGGA	GGNGGCAGA	5610 UUACCAACOS	5620 CCCACCGUUG	5630 CUGGUUGUGC	5640 UCUCCCCCC
5650 CCCAAGAAGG	5660 CCCCGACGCC	UCCCCCAAGG	AGACGCCGGA	5690 CAGUGGGUCU	5700 GAGOGAGAGC
5710 ACCAUNUCAG	AAGCCCUCCA	5730 GCAACUGGCC	5740 AUCHAGACCU	5750 UUGGCCAGCC	
5770 GKJENUGCNG	. 5780 GCUCGUCCAC	5790 GGGGGCGGGC	5800 GCCGCCGANU	6810 COGCCGUCC	S820 GACGUCCCCU
5830 GGUGAGCCGG	5840 CCCCCUCAGA	SRSO GACAGGUUCC	5860 GCCUCCUCUA	5870 UGCCCCCCCU	5880 CGAGGGGGAG
5890 CCUGGAGAUC	5900 CGGACCUGGA	5910 GUCUGAUCAG			5940 CCAGGGGGGG
5950 COCCUNGCUC	5960 CCGGUUXXGG	5970 CUCGGGGUCU	5980 UGGUCUACUU	5990 GCUCCGAGGA	GGACGAUACC
ACCGUGUGUU ACCGUGUGUU	GCCCCADGUC	AUACUCCUGG	VOCEGGGCAC	UANUANCUCC	COCOMPCCCC
6070 Фалсасская	AGUUGCCAAU	CYVCCCOOOG	AGUAACUCGC	DGUUGCGAUA	
	CAACAUCAAA	6150 GAGCIGCCUCA		ANNAGGUANC	
	DOGACGCCCA	UUAUGACUCA		ACKUCANGCU	
		CACCURGAG		AGUUGACUCC	
	AGUAUGGAUU	6330 CGGGGCCAAG		GCUUGUCCGG	
		00E6 DUDDARDDAAD			
	CCAAAAAUGA	CONGUNCACA	GUĞĞAÇÇÇĞ	CCN/AGGGGGG	
6490 GOUCGCOUCA	UCGUUUACCC	06ACCUCOGC		GCGAGXAAAU	
GACAUNACAC		00000000000000000000000000000000000000			
6610 COUGOCCAAC	6620 Goyuggagua	6630 UCUCUUGAAA	6640 GCAUGGGGGG	6650 Aaaagaagga	CCCCAUGGGU CCCCAUGGGU
0000CGUAGG	AUACCCGAUG	CUUCGACUCA	6700 ACCUUCACUG	6710 NGAGAGACAU	CAGGACCGAG 6720
6730 GAGUÇCAUAU	6740 ACCAGGCCUG	6750 CUCCCUGCCC :	6760 Baggaggccc	6770 GCACUGOCAU	6780 ACACUCGOUG
6790 ACUGAGAGAC	6800 UUUNOGUNGG	6810 AGGCCCAUG	6820 AUCAACAGCA	6830 AGGGUCAAAC	CUSCGGUUAC
6850 AGACGUUGCC	6860 DECEMBRECORDS	6870 GGUGCUAACC	6880 ACUAGCADGG	. 6890 GUANCACCAU	6900 CACAUGCUAU

# [Figure 2F]

GUGAAAGCCC	6920 UAGCGGCCUG	6930 CAAGGCUGCG	6940 GGGAUAGUUG	GCCCACAAU	6961 GCUGGUADGO
. 6970	6980	6990 CUCAGAAAGO	7000	7010	7020
7030 AGAGOXINGA	7040	7050 GACCAGGUAC	7060 UCUGCCCCUC	7070 CUGGUGAUCO	7080 CCCCAGACCC
		7110			
GAAUAUGACC	OCCAGCUAAU	AACAUCCUGD	DOCUCARAUG	UGUCUGUGGG	GUUGGGCCCG
7150 CGGGGCCGCC	7160 GCAGAUACUA	7170 CCUGACCAGA	7180 GACCCAACCA	7190 CUCCACUCGO	CCGGGCDGCC
7210	7220	7230 CCCUAUCAAU	7240	7250	7260
CCAACCAUAU	GGGUUCGCAU	7290 GGUCCUAAUG	ACACACUUCU	UCUCCAUUCU	CAUGGUCCAA
7330	7340	7350 CAACUUUGAG	7360	7370	7360
BOUCOCK	ACCAMBACCO 24.00	2410	7420	7430	2440
THRESTCORNS	CHRICCAUANU	UGAGAGGUUA	CYCGGGCODG	ACCCUUOUC	UAUGCACACA
7450 UACUCUCACO	7450 AOGAACUGAC	7410 UGAGAGGUUA 7470 GCGGGUGGGU	T480 DCAGCCCICA	7490 GAAAACUDGG	7500 GGGGCCACOC
7510	7520	7630 GGCUCGCGCA	7540	7550	7560
AAAGOGGCCG	DDOGGGGGGG	7590 NUAUCUCUUC	AAUUSGGCGG	UGAAGACCAA	CCUCANACUC
7630	7640	7650 CCUACUGGAC	7560	7670	7680
GGGGGGGACA	UUUUUCACAG	7710 CGUGUCGCUC 7770 AGGCCUCUUC	GCCCGNCCCC	GCUCAUUACU	CUUCOGCCUA
. 7750	7760	7770	7780	7790	7800
2810	7820	7830	7840	7850	7860
UAGGUACACU	CCAUAGCUAA	COSTOCCOON	บบบบบบบบบบ	COCCOCCOCC	บบบบบบบบบบ
7870 <b>UUUUUUUUU</b> UUU	7680	7890 UUUUUCCCUC	7900 UUUCUUCCCU	7910 UCUCAUCUUA	7920 UUCUACUUUC
7930	7940	-7950 AGCCCUAGUC	7960	7970	7980
7990 AUGACUGCAG	0008 00000000000	0103 DUCUGGUCUC	8020 UCUGCAGAUC	9030 8030	8040

#### [Figure 3A]

ACCCCCCCCC	DS SOSSONIAL	ACACUCOGCO	AUGAAUCACU	50 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	GAACUACUG
CUUCACGCAG	ANAGOGUCUA 80	90 GCCAUQGCBU	100 UNGUAUGNGU	GUCGUACAGO	CUCCAGGCC
130	140	150	160		1.84
100	200	. 210	220		. 24
250	250	270	280		300
				350 AUGAGCACAA	
370	380	390	400	410	420
430	440	450	460	AUUGAACAAG 470	. 480
				UNUGACUGGG 530 CAGGGGGGGCC	
550	560	570	580	590	600
UGUCAAGACC	GYCCAGACCA	GUGCCCUGAA	UGAACUGCAG	GACGAGGCAG	CGCGGCUAUC
GUGGCUGGCC 670	ACGACGGGCG 680	000000000 690		GACGUUGUCA 71.0	CUGAAGOGGG
AAGGGACKGG	CUECUAUUGG	GCGARGUGCC	GGGGCNGGAU	CUCCUGUCAU	CUCACCOUGC
UCCOGCOGAG					
790 GGCUACCUGC			•	GYGCGYGCYG	CUVCACOCCYA
850 GGAAGCCGGU				CAUCAGGGGC	
OGAACUSUUC OGAACUSUUC					
970 UGGCGAUGOC	980 UGCUUGCCGA	990 AUAUCAUGGU	0001 Döğualardı	CGCUUUUCUG	
1030	1040 CUGGGUGUGG	1050 CGGACCGCUA	1060 UCAGGACAUA	GCGUUGGCUA	1080 CCCGUGAUAU
1090 UGCUGANGAG	1100 CUUGGCGGGG	1110 AAUGGGCUGA	1120 CCGCOUCCUC	1130 GUGCUUUACG	1140 GUAUCGCCGC
1150 UCCOGAUUCG	1160 CAGCGCAUCS	1170	1180 CCOUCUUGAC	1190 GAGUUCUUCU	1200 Gaguuuaaac
1210	1220	1230	1240	1250	1250
. 1270	1280	1290	1300	1310	1320
CGUUUGUEUA	UALGUUALUU   1340	UCCACCAUAU 1350	UGCCGUCUUU 1360	DGGCAAUGUG 1 1370 -	
AACCUGGCCC	nanonnonna 1	ACGAGCAUUC	CAVGCCCCCA		

# [Figure 3B]

	· 1390	1400 GUUGAAUGUU	1410 GUGRAGGARG	1420 CAGUUCCUCE	1430 GGAAGKUUKU	144 UGAAGACAA
	TANGGUGUGU	1460 AGOGACOCUU	1470 UGCAGGCAGC	1480 GGAACCCCCC	1490 ACCUGGOGAC	150 AGGUGGCUC
	1510 GOGGCCAANA	1520 GCCACGUGUA	1.530 UARGAUACAC	2540 CUGCAZAGGO	1550 GOCACAACCO	156 CASTIGCCAC
	1570 UUGUGAGUUG	3580 GAUAGOUGUS	1590 GAAAGAGUCA	1600 NAUGGCUCUS	1610 CUCAAGOGUA	162 UUCAAÇAAG
	1630 GGCUGARGGA	1640 UGCCCAGAAG	1650 GUACCCAUU		1670 UGAUCUGGGG	CCUCGGUGC
	1690 CAUGCUUUAC	1700 AUGUGUUUAG	1710 UKKTAGOUUAA	1720 ЛАЛАЛСКИ	AGGCCCCCCC	ARCCACGGG
	1750 AGGGGGGGGG	CCUUUGABAA AAAAGUUUCO		ACCAUGGCCC		DUNOGOCCA
	1810 CAGACACGAG	GUCUCUUGGG	1830 CUCUAUAGUG	GUGAGCAUGA	OGGGGGGGUGA	
	1870 CAGGCCGGGG	1880 Aggucchagu	1890 COUGUCCACA	1900 Gucacucagu	CCUUCCUCGG	Archucchur Archucchur
	1930 UCGGGGGUCU	UAUGGACUGU	UURCCACGGA	GCUGGCAACA	AGNCACUNGO	
	GGCCCGGUCA	COCAGAUGUA	CACGYGCGCC	GAGGGGGACU	Deanceeine	
	COUGGGACCA	ANUCUUUGGA	GCCGGGGNVCG	UGUGGAGCGG	TICGACCUGUA	
		AUGUCAUCCC		CGCGGGGACY		
		2180 UUUUGACCOU		UCGGGGGGAC		
	2230 CAOGCUGUOG	GVARICADOCCE GVARICADOCCE				
	2298 UUCAUCCCCG	2300 UKRAGACGEU				
	ACACCACCAG	Q360 A3CCCCA	2370 GACCUAUCAG	2360 GUGGGGUACU	UGCACGCCCC	
	2410 GGAAAAAGCA	2420 CCANGGUÇÇÇ	2430 CGOCGCGUAC	GCCGCCCAGG	2450 GGUNUAANGU	BCORROGO 3460
	2470 ЛАИСССИОЭЭ	OASSECTION	2490 CCUGGGRUUU	GOGGCGUACU	2510 UGUCCÁAGGC	2520 ACAUGGCAUC
	2530 AACCCCAACA	254D UUAGGACUGG	2550 Agucagaacu	2560 GUGACGACO3	2570 GGGAGCCCAU	2580 UACAUACUCC
	2590 ACGUAUGGUA	2600 ANUUCCUUSC	2610 CGAUGGGGGC	2620 UGOGCAGGOG	2630 GOGCCUAUGA	2640 AUGADGADA
	2650 UGOGAUGNAU	2660 GOCACUCURU	2670 GGAUGCUACC	2680 ACUAUUCUCO	2690 GCAUCGGGAC	2700 AGUCCUUSAC
•	. 2710 CAAGCAGAGA	2720 CMGCCGGGGU	2730 CAGGCUAACU	2740 GUACUGGCCA	2750 COGCCACOCC (	2760 CCCCGGGUCG

# [Figure 3C]

GUGACAACCC	2780 COCADCOCAA	2790 UAUAGAGGAG	2800 GUNGCCCUCG	2810 GACAGGAGGG	2820 UGAGAUCCCC
· 2830 UUCUAUGGGA	2840 GGGCGUURICO	OCUGUCUUAC	2860 AUÇAAGGRAG	2870 GGAGGGACUU	2880 GAUUUUCUGO
2890 CACUCAAAGA	2900 AAAAGUSUGA	2910 CGAGCUCGCA	2920 ACGGCCCOUC	2930 GGGGCNUGGG	2940 CUUGAACGCU
2950 GUGGCAUAUU	2960 ACRGAGGGUD	2970 GGACGUÇUCC	DBES AASSAUKAUK	0999 CIXCAAGOAGA	UGURRUGGUR
3910 GUUGCCACCG	90300000000000000000000000000000000000	3030 GACGGGGGUAU	3040 ACUGGAGACU	3050 UUGACUCOGU	GAUCGACUGO
		CGUAGACOUC	AGCCUGGACC	CCACCUUCAC	UAUAACCACA
3330 CAGACUGUCE	3140 CGCAAGACGC	uguçuc#agu			
	UUUNUNGGUN	UGUUUUCCACU	GGUGAGCGAG	CCUCAGGAAU	GUUUGACAGU
		3270 CGACGCAGGA			ACCAGUGGAG
		GUAUUUCAAC	ACGCCUGGCO	OCCCOGUGUS	CCYCCYCCYC
		UUUCACCGGC			
	agucogggga.	3450 AAAUUUOGCA	DACUUAGUAG	CCUAUCAGGC	CYCYGOGG
		cccguccuss	GACGUCAUGU		GYCACAYCAC
		DACACCUCUC	COCONCCCOO	Deeccococo	DACCARCSAG
		GACAAAAUAC	3640 NUCOCCACAU 3700	GCAUGCAAGC	OGNOCUCGAG
		CCUGGCUKGG-	GGAGUCUUAG	CAGCCGUCGC	CGCGUAUUGC
	GGUGUGURRIC	CAUCAUUGGC	CGUUUACACA 3820	UCAACCAGOG 3830	AGCUGUCGUÓ 3840
		3810 CUAUGAGGCU 3870	UUUQAUGAGA 3680	UGGAGGAAUG	UGCCUCCAGA
		GCAGCGGAUA	GCCGAGAUGC	UGANGUCCAA	GAUCCAAGGC
		ACAGGOCCAG	GACAUACAAG.		
3970 CCCAAGAUGG	3980 AGCAAUTICUG	3990 GGCCAAACAU	AUGUGGAACU	4010 UCAUNAGOGG	4020 Схинсление
4030 COCGCAGGAC	4040 UGUCAACACU	4050 GCCAGGGAAC	4050 CCUGCUGUGG	4970 CUUCCAUGAU	4080 COCAUUCAGO
. 4090 GCCGCCCUCN	- 1100 - 1100	4110 GIXAACUAGC	4120 ACCACCAUCC	4130 UUCUUAACAU	6140 UCUGGGGGGC

### [Figure 3D]

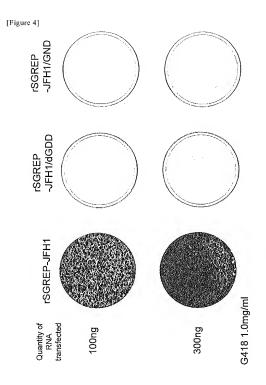
UGGCUGGCGU	41.60 CCCAAAUUGC	4170 GCCACCCGCG	6GGGCCACGG	GCUUUGUUGU GCUUUGUUGU	CAGUGGCCU
4230		4230	4240	4250	426
4276 UAUGGUGOGG	4280 GCAUUUCGGG	4290 GGCCCUCGUC	4300 GCGUUUAAGA	4310 DCAUGUCUGG	CGAGAAGCC
4330 UCCAUGGAGG	4340 AKEKAUCAA	. 4350 COUSCUGOCU	4360 030AUUCUGU	CUCCAGGUGG	4381 UCUGGUGGU
4390 GGAGUÇAUCU	4400 GOSCOGCCAU	UCUGCGCCGC	CAUGUGGGAC	CGGGGGAAGG	COCCOCICCAI
4450 UGGAUGAACA	4460 GGCUUAUCGC	auticsautice	AGAGGAAACC	ACGUCGCCCC	UACUCACUAC
	4520 COGAUGOGUO	GCAGCGUGUC	ACCCARCUGG	nnecchanan	CACUAUAACI
	4580 GGAGACUUCA				
	4640 GCGAUGUGUG	GGACUGGGUC	UGUACCAUCC	UARÇAGACUU	0580 055650 4740
	AGCUGUUCCC	4710 AAAGAUGCCU 4770	eeccaccaca	UUNUCUCUUG	CCANANGGGG
UACAAGGGCG 4810	4760 UGUGGGCCGG 4820	CACUGGCAUC 4830	AUGACCACAC	GAUGCCCCUG 4850	CGGCGCCAAC
ADCUCUGGCA 4870	ACGUCCGCUU 4860	GGGCUCUNTIG 4890	AGANUCAÇÃO 4900	GACCCAAAAC 4910	CUGCAUGAAC 4920
ACCUSSCAGE 4930	GGACCUTUICC 4940	UAUCAAUUGU 4950	uauagaag 4960	GCCAGUGCOU	GCCGAAACCC
GCGUUNANCU 4990	UCANGACOSC 5000	Caucucgágá 5010.	GUGGGGGGCCU	CAGAGDACGG 5030	GGAAGUGACG 5040
CAGCACGGAU	CAUAUGCOUA	UAUAACAGGG	CUGACCACUG	ACAACUUAAA	AGUCCCUUGC
CAACUCOCCU 5110	5060 CUCCAGAGUU 5120	5130		UACAAAÜÜÜÄ 5150	UAGGUCCGCC 5160
CCCACACCAA	AGCOGUUUUU	CCGCGAUGAG	GUCUCGUUCA	ecennegecn	CYVORYOOD
GUCGUCGGG0	CUCAGCUUCC 5240	CUGUGACCCO 5250	GAGCCCGACA 5260	CUGNGGUNGU 5270	gauguccaúg 5280
CUANCAGACC 5290	CAUCKICAUAU 5300	CACOGCOGAG 5310	CCOCCACCCC		GOGGGGÜCĂ 5340
COCCAUCUS	AGGCAAGCUC	CUCAGCGAGC	CAGCUGUCGG	CCCCAUCCCO	GCGNGCCNCC
	DADDAUDOOA				
5410 GGCGGCGUGA	5420 UUCGGAUAGA	\$430 GUCUGAGUCC	5440 NAAGUGGUCG	5450 Ducuggacuc	9460 ADUDADOUDO
5470 AUGACOGAGG	5480 AAGAGGGCGA	5490 CCTUGAGCCU	5500 UCAGNACCAU	5510 CGGAGUAUAU	5520 GCIXCCCAGG

# [Figure 3E]

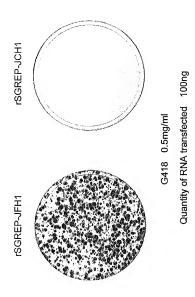
AAGAGGGGGCC	CACCOGCCUU	5550 ACCGCUUGG	9590930003	5570 AUUACAACCC	ACCGCUUGU
5590 Graucougga	5600 AGAGGCCAGA	5610 UUACCAACCA	S620 CCCACUGUUS	5630 CGGGCUGUGC	UCUCCCCCC
5650 CCCARARAGA	5660 CCCCGACGCC	UCCUCCANGG	5680 AGACGCCGGA	5690 CAGUGGGUCU	570 GAGCCAGAG
5710 ACCAUAGGAG	5720 AUGCCCUCCA	5730 ACAGGUGGCC	5740 AUCAAGUCCU	5750 OUGGCCAGCC	COCCCARG
5770 GGCGAUUCAG	5780 GCCUUUCCAC	. 5790 GGGGGCGGAC	GCCGCCGACU	COSGCGNUOG	GACACCCCC
		GACAGGUUÇU	ACCUCCUCCA	necocccccn	CSAGGGGGAG
	CAGACCUGGA	GCCUGAGCAG	GUAGAGCUUC	VFOCACCACC	CCAGGGGGGG
		cucagagaca		GCUCCGAGGA	
	GCUCCAUGUC	NUAUUCCX33	ACCOGGGCUC	DANDACUCC	TURUAGOOO
	AGUUGCCAAU	UAACUKXUUG	MGCANGUOGC	UGUUGÇGAŲA	CCAUAACAAC
GUAUACUGUA 6190	CUNCAUCAAA 6200	GASUGOCOCA 6210	CURAGGGCUA 6320		UUUUUNUNG 6246
AUGCAAGUGC	UCGACGCCUĂ. 6760	UDAUGAUUCA 6270	GUCUUAÃAGO 6280	ACAUCAAGCU 6290	AGOGGCCUCC 6300
AAGGUCAGOG 6310	CANGGOÜCCÜ 6320	CACCUUAGAG 6330	GAGGCGUGCC	ANUUGACCCC	<b>ACCOCACUCE</b>
GCAAGAUCCA.	AGUAUGGGUU 6380	DGGGGGGAAG 6390	GAGGUCUCCA G400	GCUUGUCCGG 6410	GAEGGCCGUG 6420
AACCACAUCA	AGUECGERGUG	GNAGGACCUC	UUGGAAGACU	CACAAACACC	6480 6480
		GGUGUUCUGC	CUGGACCCCG	CCANGGGGGG	UAAAAAACCA
6490 GCUCGCCUNA		DGACCIICOGC	GUCAGGGUCU	GCCAGAGAGAU	BGCCCTUUAU
6550 GAUGUCACAC	AAAAGCOOCC	UCAGGCGGUG			
6610 CCCGCXCAGC	6620 GGGUGGAGUU	6630 DCUCUUGAAG	GCAUGGGCGB	6650 Aragagaga	
6670 UUUUCGUAUG	6680 AUACCCGAUA	6690 CUUUGACUCA	6700 ACCGUCACUG	6710 Agagagacau	6720 CAGGACUGAG
6730 GAGUCCAUAU	6740 ACCAGGCCUG	6750 CUCCUVACOC	6760 GAGGNGGCCC	G770 GAACUGCCAU	6780 ACACUUGCUG
6790 ACUGAGAGAC	6800 DCUAUGUGGG	6810 AGGGCCAUG	6820 GBCAACAGCA	6830 AGGGCCAGÚC	6840 CUGCGGGUNC
6850 AGGCGUUGĆĆ	6860 BOGGCAGOGG	6870 GGUGCUUACC	6880 ACUAGUAUCG	6890 GUANCACCAU	6900 CACAUGCDAU

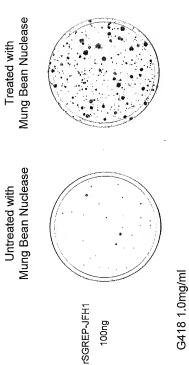
# [Figure 3F]

691.0 GUAAAAGCCC		6930 CAAGGCXGCG	6940 GGGAUAAUDG	6950 CGCCCACGAU	696 GCUGGUAUG
GGCGACGACU			7000 CAGGGGAÇUG	7010 AGGAGGACGA	702 GCGGAACCU
. 7030 AGAGCCUUCA		7050 GACCAGGUAU	7060 UCUGCCCCUC	7070 CUGGUGACCC	708 CCCCAGACC
	7100 UGGAGCOAAU		7120 UCCUCARAGE	7130 UGUCUGUGGC	ACUUGGCCC)
7150 CAGGGCGGCG		CCUGACCAGA	GACCCCACCA	CUUCAAUUGC	7200 CCGGGCUGCG
7210 UGGGAAACAG	UUAGACACUO	CCCUGUCAAU		GAAACAUCAU	
	GOZUUCGCAU	CONCORNO	ACACACUDOU	UCUCCAUKU	CAUGGCCCAG
	ACCAGAACCU	DAACUUUGAA	AUGUACGGAU	COGUGUACUC	OGUGAGUOCU
	CAGCCAUANU	UGAAAGGÜÜA	CACGGGCUUG	VOCCCOOCC	UCUGCACACA
		èceagnascu		7490 GAAAACUXSG	
				7550 CCCUCAUCIC	
				7610 DGAAGACCAA	
ACUCCOUUSC	COCAGGCVCG	CCACCACCAC	UDGUCCAGUU	7670 GGUUUACCGU 7730	00000000000
GOGGGGGACA	UUUAUCAÇAG	CGUCUCGCGu	GCCCGACCCC	GCCUNUUACU	CCUUAGCCUA
	CUGUNGGGGU	AGGCCCCCUUC	CUNCUCCCCG	7790 CUCGAUAGAG 7850	OGGCACACAU
UAGCUACACU	CCAUAGCUAA	COCOCCOOL	00000000000	7910	
UUUUUUUUUU	CINTILITIA	uuuuucceue	OOOCOOCCO 1	7970 7970	DUCOVCOTOG
UUDCUUGGUG	GCUCCAUCUU	AGOCCUAGUC	acggoungou i	SUGNAAGOUC (	CGUGAGCCGC
AUGROUGORG	AGAGUGCOGU	AACUGGUCUC I	UCUGCAGAUC I	9030 8030	- \$U40 -

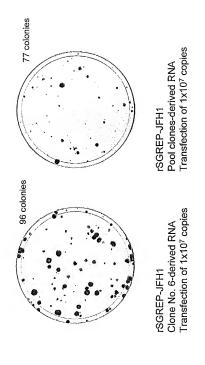


[Figure 5]

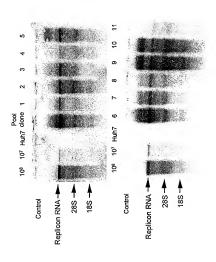




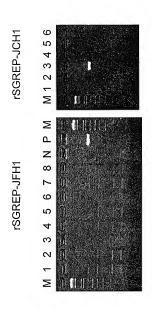
[Figure 7]



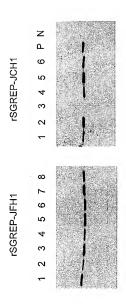
[Figure 8]



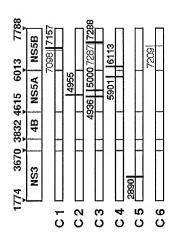
[Figure 9]



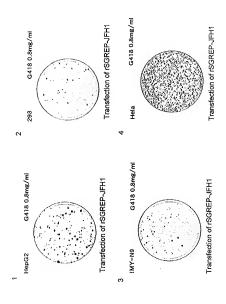
[Figure 10]



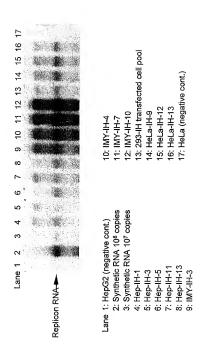
[Figure 11]



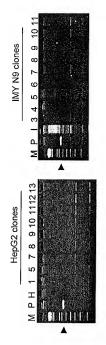
[Figure 12]



[Figure 13]

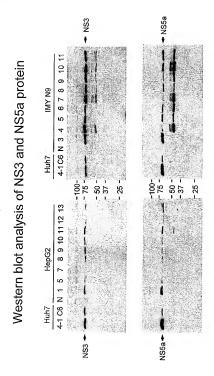


Detection of neomycine resistent gene integrations In HepG2 and IMYN9 replicon cells by genomic DNA PCR analysis



M: DNA size marker
P: Positeve control
H: HepG2 cell
I: IMYN9 cell
➤: PCR product

[Figure 15]



[Title of Document] ABSTRACT

[Abstract]

[Technical Problem] An object is to provide a replicon RNA that is derived from HCV of a different genotype from genotype 1b.

[Technical Solution] A replicon RNA comprising a nucleotide sequence at least containing the 5' untranslated region, the nucleotide sequence encoding NS3 protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein, and the 3' untranslated region on the genomic RNA of hepatitis C virus of genotype 2a is provided.

[Selected Drawing] None